

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:38:16 ; Search time 23.38 Seconds  
(without alignments)  
697.011 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSRQDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_1101.\*

- 1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1159	99.8	220	AAE04837	Human SGP003 phosph
2	1026.5	88.4	221	AAE18667	A human regulator
3	465.5	40.1	198	AAE73219	Human phosphatase
4	465.5	40.1	248	AAE68779	Amino acid sequenc
5	465.5	40.1	318	AAE42334	Human ORFX ORF2098
6	463.5	39.9	221	AAE92177	Human cardiovascular
7	461	39.7	198	AAE73214	Murine phosphatase
8	460.5	39.7	327	AAE39977	Human polypeptide
9	456.5	39.3	291	AAE28519	Human DSP-7, Homo
10	445	38.3	211	AAE5620	Human dual specifi
11	445	38.3	211	AAE39356	Human polypeptide

12	445	38.3	211	22	AAE04839	Human SGP060 phosph
13	445	38.3	328	22	AAE42355	Human polypeptide
14	443.5	38.2	549	22	AAE04838	Human SGP014 phosph
15	426	36.7	207	22	AAE28793	Human hydrolase-11
16	377	32.5	1133	22	AAE04836	Human SGP018 phosph
17	363	31.3	263	21	AAE92175	Human cardiovascular
18	346.5	29.8	180	22	AAE06787	Human dual-specifi
19	346.5	29.8	180	22	AAE66441	Human MAP-kinase p
20	346.5	29.8	185	15	AAE56968	Human phosphatase
21	346.5	29.8	185	18	AAE35330	Human cdc25B vacci
22	346.5	29.8	185	21	AAE42873	Human ORFX ORF2637
23	302	26.0	150	21	AAE40919	Human ORFX ORF683
24	288	24.8	276	22	AAE39978	Human polypeptide
25	284	24.5	240	21	AAE28520	Human DSP-7 splice
26	256	22.0	167	21	AAE5621	Human dual specifi
27	247.5	21.3	118	15	AAE56969	Human phosphatase
28	224	19.3	173	22	AAE41763	Human polypeptide
29	224	19.3	173	22	AAE41764	Human polypeptide
30	206.5	17.8	205	22	AAE81105	Anino acid sequenc
31	205	17.7	184	21	AAE68795	Human phosphatase
32	205	17.7	184	22	AAE73216	Human dual-specifi
33	205	17.7	184	22	AAE67167	Human DSP-3 protei
34	205	17.7	184	22	AAE66431	Murine phosphatase
35	200	17.2	184	22	AAE73213	Murine DSP-3 varia
36	200	17.2	205	22	AAE66443	Human SGP014 phosph
37	197	17.0	74	22	AAE04854	Human protein sequ
38	195.5	16.8	179	22	AAE25664	Human polypeptide
39	195	16.8	80	22	AAE41142	Human dual-specifi
40	192.5	16.6	169	22	AAE06783	Human MAP-kinase p
41	192.5	16.6	169	22	AAE66439	Human lung tumour
42	192.5	16.6	302	22	AAE76876	Mitogen-activated
43	192.5	16.6	392	18	AAE09039	MAP-kinase-phospha
44	192.5	16.6	394	15	AAE63601	Partial MAP kinase
45	192.5	16.6	394	16	AAE78635	

ALIGNMENTS

RESULT 1

AAE04837  
ID AAE04837 standard; Protein; 220 AA.

XX AAE04837;

DT 10-SEP-2001 (first entry)

XX Human SGP003 phosphatase polypeptide.

Human: SGP003 phosphatase polypeptide; phosphatase-related disease;  
immune-related disorder; ocular disease; organ transplant rejection;  
infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
attention disorder; cognition disorder; psychotic disorder; cytostatic;  
neurological disorder; viricide; nootropic; cerebroprotective; therapy;  
neuroprotective; antibacterial; vulnery; tranquiliser; antiasthmatic;  
hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;  
antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
MKP; migraine; chromosome CHR10.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..173

FT Domain /label= Catalytic\_domain

FT Domain 54..199

FT Domain /label= Phosphatase\_domain

FT Misc-difference 85

XX /note= "Encoded by CPG"

PN WO200146394-A2.

XX 28-JUN-2001.  
XX 21-DEC-2000; 2000WO-US34736.  
XX 21-DEC-1999; 99US-0173255.  
XX 28-DEC-1999; 99US-0175766.  
XX 25-JAN-2000; 2000US-0178078.  
XX 31-JAN-2000; 2000US-0179301.  
XX (SUGEN) SUGEN INC.  
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
XX Flanagan P;  
XX WPI: 2001-418058/44.  
XX N-PSDB; AAD09495.  
XX Novel phosphatase polypeptide useful for treating cancers,  
XX immune-related diseases and disorders, cardiovascular disease, brain or  
XX neuronal-associated diseases and metabolic disorders  
XX Claim 7; Fig 2; 186pp; English.  
XX The present invention relates to phosphatase polypeptides, nucleotide  
XX sequences encoding them, as well as various products and methods useful  
XX for the diagnosis and treatment of various phosphatase-related diseases  
XX and conditions. Substance that modulates the activity of phosphatase  
XX polypeptide is used to treat immune-related diseases and disorders,  
XX cardiovascular disease, brain or neuronal-associated diseases and  
XX metabolic disorders, including cancers of tissues, cancers of  
XX haematopoietic origin, diseases of central and peripheral nervous  
XX system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,  
XX amyotrophic lateral sclerosis, viral infections, infections caused by  
XX prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,  
XX sexual dysfunction, mood disorders, attention disorders, cognition  
XX disorders, hypotension, hypertension, psychotic disorders, neurological  
XX disorders, dyskinesias and organ transplant rejection. The present  
XX amino acid sequence is human SGP003 phosphatase polypeptide. This  
XX sequence is classified as dual specificity phosphatase (DSP) and MAP  
XX kinase phosphatase (MKP). SGP003 gene maps to chromosomal position  
XX CHR10.  
XX Sequence 220 AA;

Query Match 99.8%; Score 1159; DB 22; Length 220;  
Best Local Similarity 99.5%; Pred. No. 1.2e-127;  
Matches 219; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTSGEVKTSLKNAYSSAKRLSPKMEEGEEDYCTPGAFELERLFWKSGPOYTHVNEVWP 60  
DB 1 mtsgevktslnayssakrlspkmeegeedyctpgafelerlfwkspqythvnewwp 60  
QY 61 KLYIGDEATLDRYLQKAGFTHYVNAAGRWNVDTGPDYRDMDIQYHGEADLPTFD 120  
DB 61 klyigdeataldrylqkagftthyvnaagrwvndtgpdyrdmdiqyhgeadlptfd 120  
QY 121 LSVFFYPAAAFIDRLSDHSHKILVHCVMGSRSRATLVLMTHKMTLVDAIQOVAKN 180  
DB 121 lsvffypaaafidrlsdhshkilmhcvmggrsratlvaylmthkmtlvdaigqvkn 180  
QY 181 RCVLPNRGFLKQRLDKLVQVRRSQRQDEEDGREL 220  
DB 181 rcvlpnrgflkqrlldkvlqvrrsqrdgeeedgrel 220

RESULT 2  
ID AAB18667  
XX AAB18667 standard; Protein; 221 AA.  
XX AAB18667;  
XX

DT 22-JAN-2001 (first entry)  
XX A human regulator of intracellular phosphorylation.  
DE Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;  
XX neurological disorder; Parkinson's disease; demyelinating disease;  
KW meningitis; developmental disorder; neuromuscular disorder; cancer;  
KW myasthenia gravis; cell proliferative disorder; actinic keratosis;  
KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;  
KW autoimmune disorder; inflammatory disorder; Addison's disease;  
KW acquired immunodeficiency disease; allergy; diabetes mellitus;  
KW rheumatoid arthritis; microbial infection; trauma.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Modified-site 2  
FT /note= "potential phosphorylation site"  
FT Modified-site 9  
FT /note= "potential phosphorylation site"  
FT Modified-site 16  
FT /note= "potential phosphorylation site"  
FT Modified-site 21  
FT /note= "potential phosphorylation site"  
FT Modified-site 69  
FT /note= "potential phosphorylation site"  
FT Modified-site 170  
FT /note= "potential phosphorylation site"  
FT Modified-site 208  
FT /note= "potential phosphorylation site"  
FT Active-site 146..158  
FT /note= "tyrosine specific protein phosphatase active site"  
XX WO2000055332-A2.  
XX 21-SEP-2000.  
XX 17-MAR-2000; 2000WO-US07277.  
XX 18-MAR-1999; 99US-0125593.  
XX 20-MAY-1999; 99US-0135049.  
XX 09-JUL-1999; 99US-0143188.  
XX (INCY-) INCYTE PHARM INC.  
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
XX Lu DAM, Au-Young J;  
XX WPI: 2000-602121/57.  
XX N-PSDB; AAA75684.  
XX Novel human intracellular phosphorylation regulator polypeptides and  
XX polynucleotides for diagnosis, prevention and treatment of  
XX neurological, cell proliferative and autoimmune/inflammatory disorders  
XX Claim 1; Page 84-85; 96pp; English.  
XX The present sequence represents a human regulator of intracellular  
XX phosphorylation (HRIP). HRIP is useful for screening agonists and  
XX antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
XX are useful for treating a disease or condition associated with  
XX decreased or increased expression of functional HRIP. Diseases treated  
XX or diagnosed include neurological disorders such as stroke, Parkinson's  
XX disease, demyelinating diseases, bacterial and viral meningitis and  
XX other developmental disorders of the central nervous system,  
XX neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
XX such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
XX including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
XX bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
XX inflammatory disorder such as Addison's disease, acquired  
XX immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,

CC rheumatoid arthritis, microbial infection and trauma.

XX Sequence 221 AA;

Query Match 88.4%; Score 1026.5; DB 21; Length 221;  
Best Local Similarity 91.1%; Pred. No. 4.4e-112;  
Matches 204; Conservative 1; Mismatches 12; Indels 7; Gaps 2;

QY 1 MTSEVKTSLKNAYSSAKRLSPKMEEGEEDYCTPGAFELERLFWKSGSQYTHVNEVWP 60  
DQ 1 mtsevktslknayssakrlspkmeegeedyctpgafelerlfwksgsqythvnevwp 60  
QY 61 KLYIGDEATADLRYLQKAGFTVNLAAHGRWNVDTGPDYVRMDIOYHGVEADDLPT-- 118  
DQ 61 klyigdeatadlrylqkagftvlnaaahgrwnvdtgpd---rllprghgvprrrggpttc 117  
QY 119 --FDLSVFFYPAAAFIDRALSDHKSILVHCVMGRSRSATLVLAYLMHKDMTLVDAIQ 176  
DQ 118 pfdlsvffypaaafidralsdhksilvhcvmgrsrsatlvlaylmihkdmtilvdaiq 177  
QY 177 VAKNRCVLPNRCGLKQLRELDKQLVQORRRSQRODGEEDREL 220  
DQ 178 vaknrcvlpnrgfklqreldkqlvqorrrsqrdgeedrel 221

RESULT 3

AAB73219  
ID AAB73219 standard; Protein; 198 AA.

XX AAB73219;

DT 11-MAY-2001 (first entry)

DE Human phosphatase LOC51207\_h.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
KW schizophrenia; hamartoma.

XX Homo sapiens.

XX W0200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI: 2001-211226/21.

XX N-PSDB; AAF63571.

XX New protein phosphatase polypeptide for diagnosing and treating  
PT phosphatase related disorders such as cancer, schizophrenia, cardiac  
PT dysfunction and/or vascular disorders

XX Claim 6; Fig 5; 138pp; English.

CC The present invention relates to phosphatase proteins and coding  
CC sequences. The present sequence is one such phosphatase. Phosphatases are  
CC enzymes that catalyse the dephosphorylation of proteins modified by  
CC phosphorylation of serine, threonine or tyrosine residues. The  
CC phosphatases are useful for treating a variety of diseases: for example  
CC cancer e.g. breast, uterine, prostate, head, neck, lung cancers,  
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,  
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,

CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological  
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,  
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,  
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan  
CC Zonana syndrome, schizophrenia and hamartomas.

XX Sequence 198 AA;

Query Match 40.1%; Score 465.5; DB 22; Length 198;  
Best Local Similarity 48.9%; Pred. No. 2.6e-46;  
Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;

QY 22 PKMEEGEEDYCTPGAFELERLFW-KGSPQYTHVNEVMPKLYIGDEATADLRYLQKAG 80  
DQ 12 pkmgavqspqptptlaslqllwvraqatlnhldewvpslfigdayaardksllqig 71  
QY 81 FTHVLNAAHGRWNVDTGPDYVRMDIOYHGVEADDLPTFDLSVFFYPAAAFIDRALSDH 140  
DQ 72 ithvnaaagkfvdgtgkfyrgmsleyygieadnppfdlsvfyflpvarylraalsvpq 131  
QY 141 SKILVHCVMGRSRSATLVLAYLMHKDMTLVDAIQOAKNRCVLPNRCGLKQLRELDKQL 200  
DQ 132 grvlvhcamgvsrsatlvlafimiyenmtlveaiqtqvahrnpcnpsgflrlqvlidnrl 191  
QY 201 VQORRR 206  
DQ 192 gretgr 197

RESULT 4

AAY68779  
ID AAY68779 standard; Protein; 248 AA.

XX AAY68779;

XX 16-MAY-2000 (first entry)

XX Amino acid sequence of a human phosphorylation effector PHSP-11.

XX Human; phosphorylation effector; PHSP; proliferative disorder;  
KW immune disorder; neuronal disorder.

XX Homo sapiens.

XX Key

XX Modified-site

XX Modified-site

XX Modified-site

XX Active-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX W0200006728-A2.

XX 10-FEB-2000.

XX 28-JUL-1999;

XX 99WO-US17132.

XX 28-JUL-1998;

XX 98US-0123494.

XX 14-SEP-1998;

XX 98US-0152814.

```
PR 12-JAN-1999; 99US-0173482.
XX 12-JAN-1999; 99US-0229005.
XX (INCY-) INCYTE PHARM INC.
XX Hillman JL, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;
PI Patterson C, Bandman O, Au-Young J, Gorgone GA, Yue H, Azimzai Y;
PI Reddy R, Lu DAM, Shih LL;
XX WPI; 2000-183125/16.
DR N-PSDB; AA246148.
XX
XX New human phosphorylation effectors useful for the diagnosis, treatment
PT and prevention of proliferative, immune and neuronal disorders
XX Claim 1; Page 91; 142pp; English.
XX
XX AAY68769-95 and AAY68797-99 represent human phosphorylation effectors
CC (PHSP), designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not
CC given in the specification). The sequences were isolated from CDNA
CC libraries prepared from various human tissues. The PHSP proteins are
CC useful for the diagnosis, treatment and prevention of proliferative
CC disorders, immune disorders and neuronal disorders. The PHSP proteins
CC form pharmaceutical compositions which useful for treating or preventing
CC disorders associated with decreased PHSP expression/activity. PHSP
CC antagonists are useful for treating or preventing disorders associated
CC with increased PHSP expression/activity.
XX
XX Sequence 248 AA;
SQ
Query Match 40.1%; Score 465.5; DB 21; Length 248;
Best Local Similarity 48.9%; Pred. No. 3.7e-46;
Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;
QY 22 PKMEEEGEEDYCTPGAFELERLEFW-KGSPQYTHVNEVWPKLYIGDEATALDRYLRQAG 80
Db 62 pkhgavqaspyqptlasqrlwvrgaatlnhdevwpslfgdayaardskllqg 121
QY 81 FTHVLNAHGRWNVDTPGYRMDIQYHGVEADDLPFDLSVFYPAAPFIDRLSDDH 140
Db 122 ithvnaaagkfvdgtgakyrgmsleygyleadnfnfdlsyflpvyraalsvpq 181
QY 141 SKILVHCVMGRSRTLVLAYLMIHKDMLVDAITQVAKNRCVLPNRFGLKQLRELKQL 200
Db 182 grvlhcamgvsrsatlvlafimiyemntlveaigtqahnrncpsgffrlqvlndrl 241
QY 201 VQQRRR 206
Db 242 gretgr 247
RESULT 5
AAB42334
ID AAB42334 standard; Protein; 318 AA.
XX
XX AAB42334;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF2098 polypeptide sequence SEQ ID NO:4196.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnerable; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
```

```
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX Homo sapiens.
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
DR N-PSDB; AAC76343.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease
XX
XX Claim 11; Page 3392-3393; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotrophic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 318 AA;
SQ
Query Match 40.1%; Score 465.5; DB 21; Length 318;
Best Local Similarity 48.9%; Pred. No. 5.3e-46;
Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;
QY 22 PKMEEEGEEDYCTPGAFELERLEFW-KGSPQYTHVNEVWPKLYIGDEATALDRYLRQAG 80
Db 132 pkhgavqaspyqptlasqrlwvrgaatlnhdevwpslfgdayaardskllqg 191
QY 81 FTHVLNAHGRWNVDTPGYRMDIQYHGVEADDLPFDLSVFYPAAPFIDRLSDDH 140
Db 192 ithvnaaagkfvdgtgakyrgmsleygyleadnfnfdlsyflpvyraalsvpq 251
QY 141 SKILVHCVMGRSRTLVLAYLMIHKDMLVDAITQVAKNRCVLPNRFGLKQLRELKQL 200
Db 252 grvlhcamgvsrsatlvlafimiyemntlveaigtqahnrncpsgffrlqvlndrl 311
QY 201 VQQRRR 206
XX
```



Db 312 gretgr 317

RESULT 6

AY92177  
ID AAY92177 standard; Protein; 221 AA.

XX AC AAY92177;

XX DT 01-AUG-2000 (first entry)

XX DE Human cardiovascular system associated protein tyrosine phosphatase 4.

XX KW Cardiovascular system associated protein tyrosine phosphatase 4;

XX KW CSAPTP-4; cytosolic; immunomodulatory; antidiabetic; virucide;

XX KW hypotensive; cardiant; tyrosine phosphatase modulator.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 1..158

FT /label= "N-terminal\_domain"

FT /note= "unique"

FT Modified-site 39..44

FT /note= "N-myristoylation site"

FT Domain 68..221

FT /note= "phosphatase catalytic active domain"

FT Modified-site 164..169

FT /note= "N-myristoylation site"

FT Modified-site 181..184

FT /note= "N-glycosylation site"

FT Modified-site 183..186

FT /note= "Casein kinase II phosphorylation site"

FT Modified-site 218..220

FT /note= "Protein Kinase C phosphorylation site"

XX WO200018990-A2.

XX PN 06-APR-2000.

XX PD 30-SEP-1999; 99WO-US22924.

XX PF 30-SEP-1998; 98US-0163833.

XX PR 30-SEP-1998; 98US-0164193.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Acton S;

XX WPI: 2000-293136/25.

XX N-PSDB: AAA09031, AAA09032.

XX New cardiovascular system associated protein tyrosine phosphatases are used in the diagnosis and treatment of e.g. immune disorders, anti-proliferative disorders, metabolic disorders and cardiovascular disorders

XX Claim 8; Page 154-155; 156pp; English.

XX The CSAPTP nucleic acid and protein molecules are used to modulate regulation of cellular processes. CSAPTP nucleic acid and protein molecules and modulators of CSAPTP activity and expression can be used to treat a subject with a disorder characterized by aberrant CSAPTP expression or activity. These disorders can include an immune disorder, an anti-proliferative disorder, a proliferative disorder e.g. renal and lung carcinomas, a metabolic disorder e.g. diabetes, viral pathogenesis, a neural disorder, a cardiovascular disorder e.g. hypertension or coronary heart disease or a disorder rising from improper phosphorylation of a phosphorylated protein. Anti-CSAPTP antibodies are used to isolate CSAPTP by standard techniques, to facilitate the purification of natural and recombinantly produced CSAPTP from cells, to detect CSAPTP protein in cell supernatant for evaluating the level of CSAPTP expression and to monitor protein levels in tissue as part of a clinical testing procedure

CC to determine efficacy of a treatment procedure. CSAPTP nucleic acids, antibodies and protein molecules can be used in screening assays and in predictive medicine e.g. prognostic assays, monitoring clinical trials and pharmacogenetics.

XX SQ Sequence 221 AA;

Query Match 39.9%; Score 463.5; DB 21; Length 221;

Best Local Similarity 50.0%; Pred. No. 5.3e-46; Mismatches 55; Indels 1; Gaps 1;

Matches 90; Conservative 34; Mismatches 55; Indels 1; Gaps 1;

Qy 22 PKMEEGEEEDYCTGAFELERLFW-KGSPQVTHVNEVMPKLYIGDEATLDRYLQKAG 80

Db 35 pkihgvavspypptlasqlrlwvraqatlnhldewpsifldayaardkskllqig 94

Qy 81 FTHVLNAAHGRNVDTPDYYRDMIOYHGVADDLPTFDLSVFFYPAAAFIDRLSDDH 140

Db 95 ithvnaaagkfvdtagakfyrgmsleyvgieadnnpfdllsvyflpvaryraalsvpq 154

Qy 141 SKILVHCVMGRSRSATLVLAHMKMTLVDAIQVAKNRCVLPNRFGLKOLRELDKOL 200

Db 155 grvlvncamgvsrsatlvlafimiyenmtlvealgtvqahrncpnsqfirlqlvldnrl 214

RESULT 7

AAB73214

ID AAB73214 standard; Protein; 198 AA.

AC AAB73214;

DT 11-MAY-2001 (first entry)

XX Murine phosphatase AAL44705\_m.

Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia; cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia; congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease; Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome; schizophrrenia; hamartoma.

Mus sp.

WO200112819-A2.

22-FEB-2001.

11-AUG-2000; 2000WO-US22158.

13-AUG-1999; 99US-0149005.

(SUGE-) SUGEN INC.

Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

WPI: 2001-211226/21.

N-PSDB: AAF63566.

New protein phosphatase polypeptide for diagnosing and treating phosphatase related disorders such as cancer, schizophrrenia, cardiac dysfunction and/or vascular disorders

Claim 6; Fig 5; 138pp; English.

The present invention relates to phosphatase proteins and coding sequences. The present sequence is one such phosphatase. Phosphatases are enzymes that catalyse the dephosphorylation of proteins modified by phosphorylation of serine, threonine or tyrosine residues. The phosphatases are useful for treating a variety of diseases: for example cancer e.g. breast, urogenital, prostate, head, neck, lung cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer, glioblastoma, colorectal cancer and thyroid cancer, pathophysiological



```
XX Luche RM, Wei B;
XX WPI: 2000-679377/66.
XX N-PSDB: AAC62749.
XX Novel dual-specificity mitogen-activated protein kinase phosphatase
PT polypeptide used in assays to identify agents that modulate the
PT enzyme's activity, which are useful for treating cancer and autoimmune
PT diseases.
XX Claim 1; Fig 2A; 70pp; English.
XX The present sequence is given in a method relating to the
CC isolation of dual-specificity mitogen-activated protein kinase (MAPK)
CC phosphatase DSP-7. The antibody specific to DSP-7 and the antisense
CC polynucleotide of the nucleic acid encoding DSP-7 are useful for
CC detecting DSP-7 expression in a sample which comprises an RNA or cDNA
CC preparation. The antibody is linked to a support material and a
CC detectable marker and the amount of DSP-7 polynucleotide hybridized to
CC the antisense polynucleotide is determined using PCR or hybridisation
CC assay. The isolated DSP-7 polypeptide is useful for screening agents
CC that modulate DSP-7 activity. The identified agents are useful for
CC treating Duchenne muscular dystrophy, cancer, graft-versus-host disease,
CC autoimmune diseases, allergies, metabolic diseases, abnormal cell growth,
CC abnormal cell proliferation and cell cycle abnormalities.
XX Sequence 291 AA;

Query Match 39.3%; Score 456.5; DB 21; Length 291;
Best Local Similarity 48.4%; Pred. No. 5.3e-45;
Matches 90; Conservative 36; Mismatches 59; Indels 1; Gaps 1;

QY 22 PKMESEGEEDYCTGFAFELERLFW-KGSPQYTHVNEVWPVKLYIGDENTALDRYLQKAG 80
Db 105 pelhgavqspypqptlasilqlwvraqatlnhdevwpslfgdayaardksklqlg 164
QY 81 FTHVNAAGRWNYDTGPOYVRMDIOVHGVEADDLPTFDLSVFFYPAAAFIDRLSDH 140
Db 165 ithvnaagkfvdtagkfyrgmsleyygleadnnpffdlsvflpvaryraalsvpq 224
QY 141 SKILVHCVMGRSRTLVLAYLMIHKDMLVDAIQVAKNRCVLPNRFGLKQLRELDKQL 200
Db 225 grvlvhcamgvsratvialfmlcenmtlvealgtvqahrncipnsgflqlqldnrl 284
QY 201 VQRRR 206
Db 285 gretgr 290

RESULT 10
AAY85620
ID AAY85620 standard; Protein; 211 AA.
XX AAY85620;
XX 06-FEB-2001 (first entry)
XX Human dual specificity phosphatase-9 (DSP-9) amino acid sequence.
DE Dual specificity phosphatase-9; DSP-9; human; cytosolic; cell growth;
KW immunosuppressive; antiallergic; antiproliferative; autoimmune disease;
KW cancer; graft-versus-host disease; allergy; metabolic disorder.
XX Homo sapiens.
OS WO2000060100-A1.
XX 12-OCT-2000.
XX 07-APR-2000; 2000WO-US09321.
XX

07-APR-1999; 99US-0128203.
XX (CEPT-) CEPTYR INC.
XX Luche RM, Wei B;
XX WPI: 2000-656232/63.
XX N-PSDB: AAC61100.
XX Dual specificity phosphatase-9 which dephosphorylates activated
PT mitogen-activated protein kinase, used to identify agents that inhibit
PT DSP-9 activity and modulate cell proliferation, differentiation, and
PT survival.
XX Claim 1; Fig 2A; 66pp; English.
XX This invention relates to an isolated dual specificity phosphatase-9
CC (DSP-9) and its variant. The DSP-9 protein has the ability to
CC dephosphorylate an activated mitogen activated protein (MAP) kinase.
CC Included in the invention are an expression vector comprising a
CC polynucleotide encoding the DSP-9 protein, a host cell transformed by the
CC expression vector, and an antibody that specifically binds to DSP-9.
CC DSP-9 has cytostatic; immunosuppressive; antiallergic; and
CC antiproliferative activity. DSP-9 modulating agents are useful for
CC modulating a proliferative response, differentiation or survival of a
CC cell which displays contacting inhibition of cell growth, anchorage
CC independent growth or an altered intercellular adhesion property, in a
CC patient. DSP-9 agonists and antagonists are also useful for treating a
CC disorder associated with DSP-9 activity such as Duchenne muscular
CC dystrophy, cancer, graft-versus-host disease, autoimmune disease,
CC allergies, metabolic disease, abnormal cell growth, abnormal cell
CC proliferation and cell cycle abnormalities. The present sequence
CC represents the human DSP-9 protein.
XX Sequence 211 AA;

Query Match 38.3%; Score 445; DB 21; Length 211;
Best Local Similarity 54.3%; Pred. No. 7.4e-44;
Matches 89; Conservative 26; Mismatches 47; Indels 2; Gaps 1;

QY 39 FELERLFWKGSQYTHVNEVWPVKLYIGDENTALDRYLQKAGFTHVNAAGRWVDRGP 98
Db 46 felerllytgktacnhadewvpglylgdqmnnrrrlrgrgithvlnasharwr--gtp 103
QY 99 DYRRMDIQVHGVEADDLPTFDLSVFFYPAAAFIDRLSDHDKILVHCVMGRSRTLVL 158
Db 104 eayeglgirylgveahdspafdmshfqtadafihraispqgkglvhcavgsratv 163
QY 159 LAYLMIHKDMLVDAIQVAKNRCVLPNRFGLKQLRELDKQLVQ 202
Db 164 laylmlyhhtlveaikvkdhrgiipnrgfirlaldrirq 207

RESULT 11
AAM39356
ID AAM39356 standard; Protein; 211 AA.
XX AAM39356;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2501.
DE Human; neutrotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
OS
```













GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 17:30:50 ; Search time 1322.25 Seconds  
(without alignments)  
14535.235 Million cell updates/sec

Title: US-09-847-519A-1  
Perfect score: 1165  
Sequence: 1 ggcagtggtggtggtggg.....aaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- GenEmbl.\*
- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_om.\*
- 20: em\_or.\*
- 21: em\_ov.\*
- 22: em\_pat.\*
- 23: em\_ph.\*
- 24: em\_pl.\*
- 25: em\_ro.\*
- 26: em\_sts.\*
- 27: em\_sy.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htgo\_hum.\*
- 31: em\_htgo\_inv.\*
- 32: em\_htgo\_rod.\*
- 33: em\_htg\_hum.\*
- 34: em\_htg\_inv.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1096.6	94.1	1262	6	AX180878	AX180878 Sequence
2	589.4	50.6	159517	9	AC018511	AC018511 Homo sapi
3	208	17.9	1726	9	AY040091	AY040091 Homo sapi
4	204.2	17.5	636	6	AX180880	AX180880 Sequence
5	204.2	17.5	1665	9	BC001613	BC001613 Homo sapi
6	201	17.3	1917	6	AX180879	AX180879 Sequence
7	199.4	17.1	904	6	AX086020	AX086020 Sequence
8	197.8	17.0	937	9	AB027004	AB027004 Homo sapi
9	197.8	17.0	1053	9	BC009778	BC009778 Homo sapi
10	193.8	16.6	1243	9	BC003115	BC003115 Homo sapi
11	175	15.0	1691	6	AR113043	AR113043 Sequence
12	165.2	14.2	597	6	AX086010	AX086010 Sequence
13	165.2	14.2	686	10	AB027003	AB027003 Mus muscu
14	165.2	14.2	1094	10	AF237620	AF237620 Mus muscu
15	145.8	12.5	4360	6	AX180877	AX180877 Sequence
16	133	11.4	861	6	U19731	U19731 Sequence 1
17	133	11.4	861	6	HUMDSPHS	L05147 Human dual
18	133	11.4	1930	9	BC002682	BC002682 Homo sapi
19	129.4	11.1	1080	10	AF280809	AF280809 Mus muscu
20	119.4	10.2	240	6	AX180923	AX180923 Sequence
21	103.4	8.9	177597	2	AC013603	AC013603 Homo sapi
22	103.4	8.9	184022	2	AC015779	AC015779 Homo sapi
23	98	8.4	205347	2	AL392111	AL392111 Homo sapi
24	97.8	8.4	148849	9	AL158837	AL158837 Human DNA
25	96.2	8.3	165245	2	AL451050	AL451050 Homo sapi
26	78.6	6.7	4162	2	AC012984	AC012984 Drosophil
27	78.6	6.7	165097	3	AC018490	AC018490 Drosophil
28	78.6	6.7	298705	3	AE003510	AE003510 Drosophil
29	73.4	6.3	1467	5	AF167296	AF167296 Gallus ga
30	70.6	6.1	1236	9	S80632	S80632 threonine,
31	70.6	6.1	1238	6	AR075981	AR075981 Sequence
32	70.6	6.1	1238	6	AX063176	AX063176 Sequence
33	70.6	6.1	1849	6	BC002671	BC002671 Homo sapi
34	70.6	6.1	2064	9	AX063198	AX063198 Sequence
35	70.6	6.1	2109	6	AX063199	AX063199 Sequence
36	70.6	6.1	2234	9	HSU21108	U21108 Human dual
37	70.6	6.1	2234	11	G28599	G28599 human SFS S
38	70.6	6.1	2981	9	HSM802175	AL137704 Homo sapi
39	69	5.9	1619	6	AX063174	AX063174 Sequence
40	69	5.9	2241	9	HSU48807	U48807 Human MAP k
41	69	5.9	4637	6	AX063177	AX063177 Sequence
42	67.4	5.8	2436	10	AF013144	AF013144 Rattus no
43	66.4	5.7	94752	9	AC003098	AC003098 Homo sapi
44	65.8	5.6	69061	2	AC012986	AC012986 Drosophil
45	65.8	5.6	168469	3	AC007886	AC007886 Drosophil

ALIGNMENTS

RESULT 1

LOCUS	AX180878	1362 bp	DNA	PAT	06-AUG-2001
DEFINITION	Sequence 5 from Patent WO0146394.				
ACCESSION	AX180878				
VERSION	AX180878.1	GI:15132706			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 1262)				
JOURNAL	Plowman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S., Hill,R.J. and Flanagan,P.				
FEATURES	Mammalian protein phosphatases				
source	Patent: WO 0146394-A 5 28-JUN-2001;				
	Sugen, Inc. (US)				
	Location/Qualifiers				
	1..1262				
	/organism="Homo sapiens"				

$$\begin{array}{r} 960 \\ 20 \overline{) 1920} \\ \underline{180} \phantom{0} \\ 120 \phantom{0} \\ \underline{100} \phantom{0} \\ 20 \phantom{0} \\ \underline{20} \phantom{0} \\ 0 \phantom{0} \end{array}$$

1;

$$\frac{12,991}{1,871}$$



```
FEATURES             Location/Qualifiers
  source              1..636
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT          123 a   207 c   182 g   124 t
ORIGIN

Query Match          17.5%; Score 204.2; DB 6; Length 636;
Best Local Similarity 64.8%; Pred. No. 2.1e-28;
Matches 321; Conservative 0; Mismatches 168; Indels 6; Gaps 1;

Qy 231 gcctttgagctggagcgctctcttctggaaggcgagctccagtcacccagctcaacgag 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 GTCITCGAGTTGGAGCGGCTCTCTACACAGCGAAGACAGCGCTGTAAACCTGCGACGAG 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 291 gtcctggcccaagctctacattggcgatgagcgagcgctggcgctggagccgctatagctgag 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 GTCGTGGCCAGGCTCTATCTCGGACACGAGCATGGCTAAACAACCGCGGGAGCTTCGC 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 351 aaggcgggttcacgacgctgtaagcgggcccacggcgctggaacgtggacactgag 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 CGCTTGGGCATCAGCAGCTGCTCAATGCTTCACACAGCGGCTGGCGAG-----GCACG 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 411 cccgactactaccgacatggacatcagctaccacggcgctggagccgacgacactgccc 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 CCCGAGGCTATGAGGGGCTGGGCATCGCTACCTGGGTGTGAGGCCACGACTCGCCA 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 471 accttgacctgaigtgtctttctacccggcgagcctcttcacgacagcgctaaagc 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 GCGTTTGACATGAGCATCCACTTCCAGACGGCTGCCGACTTCATCCACCGGGCGCTGAGC 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 531 gacgaccacagtaagatcgtgttcactgctcactgagccgagccgagccgagccacccgt 590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 CAGCAGAGGAGAGATCCTGTCATGTGCTGTGGCGTGTGAGCCGATCCGCCACCCCTG 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 591 gtcctggcctacctgatgatacacaagacatgacccctggtggagcccatccagcaagt 650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 GTACTGGCTACCTACCTGCTGACACACACCTTACCTCTGTTGAGGGCCATCAAGAAAGTC 546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 651 gccaaagaacctgctgctctcccaacgggggtttttaaagcagctccggagctggac 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 AAAGACACACGAGGATCATCCCAACCGGGGCTTCTGTAGGCAGCTCTGGCCCTGGAC 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 711 aagcagctggtgag 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 CGCAGGCTGGCGCAG 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
BC001613 1665 bp mRNA PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:1136 IMAGE:3535215, mRNA, complete cds.
ACCESSION BC001613
VERSION BC001613.1 GI:12804418
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
M1 (bases 1 to 1665)
Strausberg, R.
Direct Submission
Submitted (03-JAN-2001), National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: qcaps-r@mail.nih.gov
Tissue Procurement: DCTD/dnp
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhur, Parvaneh Saedi, Jacqueline
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: 0 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6807668.

FEATURES             Location/Qualifiers
  source              1..1665
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="MGC:1136 IMAGE:3535215"
                    /tissue_type="Lung, small cell carcinoma"
                    /clone_lib="NIH_MGC_7"
                    /lab_host="DH10B-R"
                    /note="Vector: pOTB7"
                    334...969
                    /codon_start=1
                    /product="Unknown (protein for MGC:1136)"
                    /protein_id="AAH01613.1"
                    /db_xref="GI:12804418"
                    /translation="MCPGNWLASMTFMARFSSRSPVTRGTLEEMPTVOHPFLN
VFLERLLYTGKTAACHADWVPLVGLGDQDMANNRRLRLGITHVNLASHNRGRT
PEAYELGLIRLVGEVDSAPFMSIHQAADFTHRLSPQGGKILVHCAGVGSRA
TLVATYLMELYHLLTVEAIIKKVDRHGIIPNRGFLRLLALDRRLRQGLEA"
BASE COUNT          372 a   480 c   499 g   314 t
ORIGIN

Query Match          17.5%; Score 204.2; DB 9; Length 1665;
Best Local Similarity 64.8%; Pred. No. 1.8e-28;
Matches 321; Conservative 0; Mismatches 168; Indels 6; Gaps 1;

Qy 231 gcctttgagctggagcgctctcttctggaaggcgagctccagtcacccagctcaacgag 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 GTCITCGAGTTGGAGCGGCTCTCTACACAGCGAAGACAGCGCTGTAAACCTGCGACGAG 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 291 gtcctggcccaagctctacattggcgatgagcgagcgctggcgctgagcctatagctgag 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 GTCGTGGCCAGGCTCTATCTCGGAGACGAGCATGGCTAAACAACCGCGGGAGCTTCGC 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 351 aaggcgggttcacgacgctgctgaaacggcgccacggcgctggaacgtggacactgag 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 586 CGCCTGGGCATCAGCAGCTGCTCAATGCCCTCACACAGCGGCTGGCGAG-----GCACG 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 411 cccgactactaccgacatggacatcagctaccacggcgctggagccgacgacactgccc 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 640 CCCGAGGCTATGAGGGGCTGGGCATCGCTACCTGGGTGTGAGGCCACGACTCGCCA 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 471 accttgacctgaigtgtctttctacccggcgagcctcttcacgacagcgctaaagc 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 700 GCGTTTGACATGAGCATCCACTTCCAGACGGCTGCCGACTTCATCCACCGGGCGCTGAGC 759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 531 gacgaccacagtaagatcgtgttcactgctcactgagccgagccgagccgagccacccgt 590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 CAGCAGAGGAGAGATCCTGTCATGTGCTGTGGCGTGTGAGCCGATCCGCCACCCCTG 819
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 591 gtcctggcctacctgatgatacacaagacatgacccctggtggagcccatccagcaagt 650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 820 GTACTGGCTACCTACCTGCTGACACACACCTTACCTCTGTTGAGGGCCATCAAGAAAGTC 879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 651 gccaaagaacctgctgctctcccaacgggggtttttaaagcagctccggagctggac 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 880 AAGACACCGAGGCGATCATCCCAACCGGGGCTTCCTGAGGAGCTCTCTGGCCCTGGAC 939

QY 711 aagcagctgtgtgcag 725  
|||||

Db 940 CGCAGGCTGGCGCAG 954

RESULT 6  
LOCUS AX180879 1917 bp DNA PAT 06-AUG-2001  
DEFINITION Sequence 6 from Patent WO0146394.  
ACCESSION AX180879  
VERSION AX180879.1 GI:15132707  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1917)  
TITLE Plozman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S.,  
Hill,R.J. and Flanagan,P.  
JOURNAL Mammalian protein phosphatases  
Patent: WO 0146394-A 6 28-JUN-2001;  
Sugen, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..1917  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 368 a 596 c 565 g 388 t  
ORIGIN

Query Match 17.3%; Score 201.; DB 6; Length 1917;  
Best Local Similarity 63.0%; Pred. No. 6.9e-28;  
Matches 328; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 192 gagggaggggagggagggagggactactgcacccctggagcctttgagctggagcggtc 251  
|||||

Db 52 GAGCTGGGGGAGAGGACAAAGCCACGCTTGGCCCGAGCATCTGGAGCTGGAGGAGCTC 111  
|||||

QY 252 ttctgaaggagagtcacccagtcacccagtcacccagtcacccagtcacccagtcacatt 311  
|||||

Db 112 CTGCGGGGAGGGAAGTCTTCTGACGCGGTGGAGCAAGTTTGGCCCAACCTTTTCATA 171  
|||||

QY 312 ggcgatgagcgagcgctggaccgctataggctcagaagcggggttcacgcacgtg 371  
|||||

Db 172 GGAGATGCGGCGCACGCAACACCGCTTTGACCTGTGGAGCTGGGCATCACCCACGTG 231  
|||||

QY 372 ctgaacgcgccacgcccgtggaacgtggacactgggcccgaactactaccgcgacatg 431  
|||||

Db 232 CTGAACGCGCCCAAGGGCCCTCTACTGTACAGGCGGCGCTTCTACGTCAGCGCAGCAGT 291  
|||||

QY 432 gacatcagtcacacggctggggcgacagacactccacccttcgacctcagtgcttc 491  
|||||

Db 292 GTGAGC---TACCTGGGGGTGGCAGCCACGACCTCCCTGATTTTGACATCATGTGCTAC 348  
|||||

QY 492 ttctaccgcgcgacgttcctcatcgacagcgctgaagcagacacacagtaagatcctg 551  
|||||

Db 349 TTCTCTCTGCGGCTGACTTCATCCACCGTGGCCCTCAACACGCGCTGGGGCCCAAGGCTCTG 408  
|||||

QY 552 gttcactgctcagtcagggcgagcggtcagccacccctggtcctggcctacctgatgc 611  
|||||

Db 409 GTGCACTGTGTGGCGGTGAGCGGCTCTGCCACGCTGGTCTCGCTTACCTCATGCTG 468  
|||||

QY 612 cacaagacatgacctgtgtgagccatccacagtaagtggccaaagacccgtcgtcctc 671  
|||||

Db 469 CACCAGCGGTGTCTCTGCGCCAGGGGGTGTATCACCCTGAGGAGCAGCAGGCTGGTCTTC 528  
|||||

QY 672 ccgaacgggggcttttgaagcagctccggagagtgagcaa 712  
|||||

Db 529 CCCAACCGAGGCTTCTGTGACACGAGTCTGCAGGCTGGACCA 569  
|||||

RESULT 7  
LOCUS AX086020 904 bp DNA PAT 09-MAR-2001  
DEFINITION Sequence 17 from Patent WO0112819.  
ACCESSION AX086020  
VERSION AX086020.1 GI:13275841  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 904)  
TITLE Plozman,G.D., Martinez,R., Whyte,D., Hill,R., Flanagan,P. and  
Lioubin,M.  
JOURNAL Protein phosphatases and diagnosis and treatment of  
phosphatase-related disorders  
Patent: WO 0112819-A 17 22-FEB-2001;  
Sugen, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..904  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 183 a 282 c 249 g 190 t  
ORIGIN

Query Match 17.1%; Score 199.4; DB 6; Length 904;  
Best Local Similarity 64.3%; Pred. No. 1.6e-27;  
Matches 299; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 276 acccagctcaacgaggtctgtgcccagctctacattgctgagcgagcgcgctggac 335  
|||||

Db 200 AACCATATCGATGAGGTCTGGCCAGCCTCTTCTGGGAGATGCTACGAGCCCGGAC 259  
|||||

QY 336 cgtataggctgcagaagcggggttcacgcacgctggaacgogggccacgcccgtg 395  
|||||

Db 260 AAGAGCAAGCTGATCCAGCTGGGAATCACCCAGCTTGTGAATGCCGTGCGAGCAAGTTC 319  
|||||

QY 396 aagctgagcactggccgagctactaccgcgacatgacatccagtcacgagcggtgag 455  
|||||

Db 320 CAGGTGGACACAGGTGCCAAATTTCTACCGTGGAAATGTCCTGGAGTACTATGGCATTTGAG 379  
|||||

QY 456 gccagcagacctgcccacttcgacctcagctgtctcttctaccggcgagcgccttcac 515  
|||||

Db 380 GCGAGCAGCAACCCCTTCTTTCAGCTCAGTCTACTTTCTGCTGTGCTCGATACATC 439  
|||||

QY 516 gacagcgctgaagcgacgacacacagtaagatcctggttcactgctgctatgagcgagc 575  
|||||

Db 440 CGAGCTGCCCTCAGTGTTCGCCAAGCGCGCTGCTGTACACTGTGCGATGGGGTAAAG 499  
|||||

QY 576 cggtcagcaccctgtctgctgacctacgtatgatccaaagacatgacctggtggac 635  
|||||

Db 500 CGCTGTGCCACACTGTGCTGGCCTTCTCATGATCTATGAGAACATGACGCTGGTAGAG 559  
|||||

QY 636 gccatccagcaagtggccaaagacgctgctctcccgacccgggctcttttgaagcag 695  
|||||

Db 560 GCATCCAGCGGTGCGAGGCCACCGCAATATGCGCCTAACTCAGGCTTCTCCCGCAG 619  
|||||

QY 696 ctccgggagctggaacagcagctggtgagcagagcgagcggttc 740  
|||||

Db 620 CTCAGGTTCTGGACACACCGACTGGGGCGGAGACGGGGCGGTTC 664  
|||||

RESULT 8  
LOCUS AB027004 937 bp mRNA PRI 13-JAN-2000  
DEFINITION Homo sapiens mRNA for protein phosphatase, complete cds.  
ACCESSION AB027004  
VERSION AB027004.1 GI:6692781  
KEYWORDS protein phosphatase.  
SOURCE Homo sapiens cDNA to mRNA.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;











BASE COUNT 144 a 205 c 182 g 155 t  
ORIGIN

Query Match 14.2%; Score 165.2; DB 10; Length 686;

Best Local Similarity 57.9%; Pred. No. 3.9e-21;

Matches 293; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

```
QY 231 gctttagctgagcgctcttctggaaggcgagtcctccagtcacacccagtcacag 290
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 GCCTCTCTGAGCGATTGCTGTGGGTCGCGGACTGCCACACTGACCCACATCAATGAG 144
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 gtctggcccaagctctacattggcgatgagcgagcgcgctggaccgctatagctgcag 350
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 GTCTGGCCCAACCTTTCTTGGGAGATGCGTATCTGCCAGAGACAAGGGTCGTCTAATC 204
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 351 aagcggggttcacgcacgtgctgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 410
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 205 CAGCTGGGCATTACCCCATGTTGTAATGTGGCTGCGGCGCAAGTTCCAGGTGGACACAGGT 264
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 ccgaactactaccgcgacatggacatccagtcacacgcgcgcgcgcgcgcgcgcgcgcgc 470
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 265 GCGAAGTCTTACCGTGGAGAACCTCTGGAGTACTATGGCATTTGAGGCTGATGACAAACCCC 324
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 accttcagctcagtgctcttcttaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 530
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 325 TTCTTTGACCTCAGCGTCCACTTTCTGCTGTTCTGTTACATCAGAGATGCCCTCAAT 384
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 gacgacacagtaagaacctggttctactgcgtcatggcgccgcgcgcgcgcgcgcgcgcgc 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 ATTCCCGGAAAGCGAGTGTGCTGCTACCTGCGCTATGGGGGTGAGTCTGTCGCACAAAT 444
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 gtctgcctaccctgatgatccacagacatgacacctggtgagcgcgcgcgcgcgcgcgcgcgc 650
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 GTCTTGGCCCTTCCATGATCTTCGAGAACATGACATGATGATGATGATGATGATGATGATG 504
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 gccaagaacgcgtgcctctccggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 710
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 505 CAGGCCACCGAGATATCTGCCCACTCAGGCTTCTCCGACAGCTCCAGGTCTCGAC 564
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 aagcagctggtgacgagagcgacg 736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 565 AACAGGCTGAGGGCGGAAACAGGAAG 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 14

AF237620 1094 bp mRNA ROD 02-MAR-2001  
LOCUS Mus musculus dual-specificity phosphatase TS-DSP6 mRNA, complete cds.

ACCESSION AF237620.1 GI:13183066

VERSION 2

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1094)  
AUTHORS Aoyama,K., Matsuda,T. and Aoki,N.  
TITLE Molecular cloning of a novel dual specificity phosphatase  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1094)  
AUTHORS Aoyama,K., Matsuda,T. and Aoki,N.  
TITLE Direct Submission  
JOURNAL Submitted (19-FEB-2000) Applied Molecular Biosciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

FEATURES

source 1..1094

/organism="Mus musculus"

/db\_xref="taxon:10090"

/codon\_start=1

/product="dual-specificity phosphatase TS-DSP6"

/protein\_id="AAK15037.1"

/db\_xref="GI:13183067"

/translation="MDSLQKQLRRPKIHGAQVQSPQPTLASLQRLWVRRATLIT  
HINEVNPFLGDAYAARDKRLIQGITHVNVNAAKQFQVDTGAKFYGTPTLEYGI  
EADDNPFDFLSVHFLPVARYIRDALNIPRSVLVHCAGVRSATIVLAFMLIFENMT  
LVDAIQTVAHRDICPNSGFLRLQVLNRLRRETGL"

BASE COUNT 287 a 293 c 284 g 227 t 3 others

## ORIGIN

Query Match 14.2%; Score 165.2; DB 10; Length 1094;  
Best Local Similarity 57.9%; Pred. No. 3.5e-21;  
Matches 293; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

```
QY 231 gctttagctgagcgctcttctggaaggcgagtcctccagtcacacccagtcacag 290
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 GCCTCTCTGAGCGATTGCTGTGGGTCGCGGACTGCCACACTGACCCACATCAATGAG 346
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 291 gtctggcccaagctctacattggcgatgagcgagcgcgctggaccgctatagctgcag 350
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 347 GTCTGGCCCAACCTTTCTTGGGAGATGCGTATCTGCCAGAGACAAGGGTCGTCTAATC 406
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 351 aagcggggttcacgcacgtgctgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 410
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 407 CAGCTGGGCATTACCCCATGTTGTAATGTGGCTGCGGCGCAAGTTCCAGGTGGACACAGGT 466
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 ccgaactactaccgcgacatggacatccagtcacacgcgcgcgcgcgcgcgcgcgcgcgc 470
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 467 GCGAAGTCTTACCGTGGAGAACCTCTGGAGTACTATGGCATTTGAGGCTGATGACAAACCCC 526
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 accttcagctcagtgctcttcttaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 530
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 527 TTCTTTGACCTCAGCGTCCACTTTCTGCTGTTCTGTTACATCAGAGATGCCCTCAAT 586
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 gacgacacagtaagaacctggttctactgcgtcatggcgccgcgcgcgcgcgcgcgcgcgc 590
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 587 ATTCCCGGAAAGCGAGTGTGCTGCTACCTGCGCTATGGGGGTGAGTCTGTCGCACAAAT 646
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 gtctgcctaccctgatgatccacagacatgacacctggtgagcgcgcgcgcgcgcgcgcgcgc 650
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 647 GTCTTGGCCCTTCCATGATCTTCGAGAACATGACATGATGATGATGATGATGATGATGATG 706
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 gccaagaacgcgtgcctctccggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 710
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 707 CAGGCCACCGAGATATCTGCCCACTCAGGCTTCTCCGACAGCTCCAGGTCTCGAC 766
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 aagcagctggtgacgagagcgacg 736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 767 AACAGGCTGAGGGCGGAAACAGGAAG 792
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 15

AX180877 4360 bp DNA PAT 06-AUG-2001  
LOCUS Sequence 4 from Patent WO0146394.

ACCESSION AX180877

VERSION AX180877.1 GI:15132705

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4360)

AUTHORS Plowman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S., Hill,R.J. and Flanagan,P.

TITLE Mammalian protein phosphatases

JOURNAL Patent: WO 0146394-A 4 28-JUN-2001;

FEATURES

source 1..4360

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 1138 a. 1076 c 1363 g 783 t

Search completed: February 8, 2002, 18:42:47  
Job time: 4317 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:40:01 ; Search time 12.62 Seconds  
(without alignments)  
392.292 Million cell updates/sec

Title: US-09-847-519a-2  
Perfect score: 1161  
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRDGEEDGREL 220

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pap.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pap.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pap.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pap.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pap.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	465.5	40.1	198	4	US-09-163-833-2
2	426	36.7	207	4	US-09-013-881-4
3	363	31.3	263	4	US-09-164-193-5
4	346.5	29.8	185	1	US-07-988-273-2
5	346.5	29.8	185	3	US-08-848-810-25
6	346.5	29.8	185	4	US-09-164-193-21
7	346.5	29.8	185	5	PCT-US93-12019-2
8	247.5	21.3	118	1	US-07-988-273-3
9	247.5	21.3	118	5	PCT-US93-12019-3
10	192.5	16.6	393	2	US-08-990-379-4
11	192.5	16.6	394	2	US-08-530-290-23
12	182.5	16.6	395	2	US-08-990-379-5
13	188.5	16.2	313	2	US-08-990-379-7
14	184.5	15.9	314	4	US-09-164-193-22
15	177.5	15.3	198	4	US-09-045-973-5
16	169	14.6	367	2	US-08-530-290-24
17	168	14.5	226	4	US-09-045-973-8
18	166	14.3	367	2	US-08-990-379-6
19	157.5	13.6	397	2	US-08-990-379-8
20	142.5	12.3	117	1	US-07-988-273-4
21	142.5	12.3	117	5	PCT-US93-12019-4
22	128.5	11.1	353	4	US-09-013-881-3
23	123	10.6	72	2	US-08-530-290-20
24	113	9.7	72	2	US-08-530-290-19
25	102.5	8.8	201	2	US-08-933-750C-21
26	102.5	8.8	201	4	US-09-234-613-21
27	100.5	8.7	116	1	US-07-988-273-5

28	100.5	8.7	116	5	PCT-US93-12019-5
29	97.5	8.4	1442	1	US-08-015-986A-3
30	97.5	8.4	1442	2	US-08-446-383-3
31	93.5	8.1	245	1	US-08-015-985-8
32	93.5	8.1	263	2	US-08-685-992-5
33	93.5	8.1	263	2	US-09-144-925-5
34	93.5	8.1	1445	1	US-08-015-986A-2
35	93.5	8.1	1445	2	US-08-446-363-2
36	85	7.3	173	3	US-08-725-532A-6
37	85	7.3	173	4	US-09-164-193-20
38	84	7.2	173	4	US-09-164-193-2
39	83.5	7.2	150	4	US-09-164-193-23
40	80.5	6.9	563	4	US-08-987-691A-2
41	79.5	6.8	242	1	US-08-015-985-7
42	79.5	6.8	360	2	US-08-531-927B-2
43	79.5	6.8	360	4	US-09-041-886-13
44	79.5	6.8	2308	1	US-08-015-973-1
45	79.5	6.8	2308	2	US-08-448-164-1

ALIGNMENTS

RESULT 1  
US-09-163-833-2  
; Sequence 2, Application US/09163833  
; Patent No. 6268135  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NO. 6268135el Phospholipase Molecule and Uses Therefor  
; FILE REFERENCE: mnl-059  
; CURRENT APPLICATION NUMBER: US/09/163,833  
; CURRENT FILING DATE: 1998-09-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 198  
; TYPE: PFT  
; ORGANISM: Homo sapiens  
US-09-163-833-2

Query Match 40.1%; Score 465.5; DB 4; Length 198;  
Best Local Similarity 48.9%; Pred No. 2.8e-46;  
Matches 91; Conservative 36; Mismatches 58; Indels 1; Gaps 1;

QY	22	PMEEEGEEEDYCTPGAFELERLFW-KGSPQYTHVNEVWPKLYIGDEATDRLQKAG	80
Db	12	PKIHGAVQASPYQPTLASLQRLWVQAATLNHIDEVWPSLFLGDAYAARDKSKLIQLG	71
QY	81	FTHVLNAAGRNVDTPDYRDMDIQYHGVADDLPTFDLSVFFYPAAAFIDRLSDDH	140
Db	72	ITHVYNAAGKQVDTPGAKFYRGSLEYGIEADDPNPFDLVYFLPVARYIRAAALSVPQ	131
QY	141	SKILVHCVMGSRATLVLAYLMIHDKMTLVDAIOQVAKNRCVLPNRFGLKOLRELDKOL	200
Db	132	GYLVHCAGVSRATLVLAFELMYENMTLVFAIQVQAHNRICPNSGFLRQLQVLDNRL	191
QY	201	VOQR 206	
Db	192	GRETGR 197	

RESULT 2  
US-09-013-881-4  
; Sequence 4, Application US/09013881  
; Patent No. 6132964  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.

APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/013,881  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: BILLINGS, LUCY J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0470 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: MUSCNOT02  
CLONE: 971204  
US-09-013-881-4

Query Match 36.7%; Score 426; DB 4; Length 207;  
Best Local Similarity 53.0%; Pred. No. 1.2e-41;  
Matches 87; Conservative 26; Mismatches 45; Indels 6; Gaps 2;

QY 39 FELERLFWKGSPOYTHVNEVWPGLYIGDEATLDRLYRLOKAGFTHVLAHGRNVDGTP 98  
DB 46 FELERLYTGTACNHADEVWPGYLGQDMANNRRLRLGITHVLNASHSRWR--GTP 103  
QY 99 DYRMDIQYHGVEADDLPTFDSVFFYPAAAFIDRLSDHSLVHCVMGSRSRATLV 158  
DB 104 EAYEGLGIRYLVE---PAFDMSTHFQTAADFIRHLSQPGGKILVHCAVGSRSATLV 159  
QY 159 LAYLMHKDMTLVDIAIQVAKNRCVLPNRFGLKQLRELDKQLV 202  
DB 160 LAYLMYHHLTLVEAIKVKDHRGIIPNRFGLQLALDRRLQ 203

RESULT 3  
US-09-164-193-5  
Sequence 5, Application US/09164193C  
Patent No. 6258582  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan L.

TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)  
FILE REFERENCE: MNI-051  
CURRENT APPLICATION NUMBER: US/09/164,193C  
CURRENT FILING DATE: 1998-09-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 263

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-164-193-5

Query Match 31.3%; Score 363; DB 4; Length 263;  
Best Local Similarity 51.8%; Pred. No. 3.6e-34;  
Matches 73; Conservative 22; Mismatches 44; Indels 2; Gaps 1;

QY 41 LERLEWKGSPQYTHVNEVWPGLYIGDEATLDRLYRLOKAGFTHVLAHGRNVDGTPDY 100  
DB 5 VRRLLYTGTACNHADEVWPGYLGQDMANNRRLRLGITHVLNASHSRWR--GTP 62  
QY 101 YRMDIQYHGVEADDLPTFDSVFFYPAAAFIDRLSDHSLVHCVMGSRSRATLVLA 160  
DB 63 YEGLGIRYLVEAHDSAPFDMSTHFQTAADFIRHLSQPGGKILVHCAVGSRSATLVLA 122  
QY 161 YLMHKDMTLVDIAIQVAKNR 181  
DB 123 YLMYHHLTLVEAIKVKDHR 143

## RESULT 4

US-07-988-273-2  
Sequence 2, Application US/07988273  
Patent No. 5512434  
GENERAL INFORMATION:

APPLICANT: AARONSON, Stuart A.  
APPLICANT: BOTTARO, Donald P.  
APPLICANT: ISHIBASHI, Toshio  
APPLICANT: MIKI, Toru  
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN  
TITLE OF INVENTION: PHOSPHATASE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,273  
FILING DATE: 19921214  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/182 NIHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 185 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-273-2

Query Match 29.8%; Score 346.5; DB 1; Length 185;  
Best Local Similarity 43.8%; Pred. No. 1.8e-32;  
Matches 78; Conservative 35; Mismatches 54; Indels 11; Gaps 3;

QY 37 GAFE-----LERLFWKGSPOYT---HVNVEVWPGLYIGDEATLDRLYRLOKAGFTHVLA 87  
DB 3 GSFELSVQDLNDLLSDGSGCYSLPQPCNEVTPRIYVGNASVAQDIPKQLKLGITHVLNA 62



Db

63 AGRSPMHVNTNANFYKDSITYLGIKANDTQEFNL SAVFERAADFDQAALAKNGRVLV 122

QY

146 HCVMGSRSRATLVAYLMIHKDWTLVDATIQOVAKNRCVLPNRGFLKQLRELDKQLVQQ 203  
|| ||| ||| : : : : : ||| : : : : : ||| : : : : :

Db

123 HCREGYRSPTLVIAIYLMRMQRMDKVSASIVRQNREIGPNDCGFLAQCLQNDRLAKE 180

```

RESULT      8
US-07-988-273-3
; Sequence 3, Application US/07988273
; Patent No. 5512434
;
; GENERAL INFORMATION:
; APPLICANT: AARONSON, Stuart A.
; APPLICANT: BOTTARO, Donald P.
; APPLICANT: ISHIBASHI, Toshio
; APPLICANT: MIKI, Toru
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

[illegible]

```

RESULT          9
PCT-US93-12019-3
; Sequence 3, Application PC/TUS9312019
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12019
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,273
FILING DATE: 14-DEC-1992
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US93-12019-3

```

	Query Match	21.3%	Score 247.5;	DB 5;	Length 118;
	Best Local Similarity	47.0%;	Pred. No. 3e-21;		
	Matches 55;	Conservative 24;	Mismatches 35;	Indels 3;	Gaps 2;
Qy	84	VLNAAGHR--WNVDTGPDYRMDIQYHGVEADDLPTEDSLVSFFYPAA-AFIDRAISDDH	140		
Db	1	VLNAACGRSFMHVTNANFYKDSGITYLGKANDTQEFNL SAYFERAADIFIDQALQAKN	60		
Qy	141	SKTLVHCVMGRSBSATLVLAYLMHKDMTLVDIAIQOVAKNRCVLPNRCFTLKQLBELD	197		
Db	61	GRYLTVHCREGYSRSPITLVATLMMRQKMDVKSALSVIRQNREIGDPNDGFLAQJCOLN	117		

```

RESULT 10
US-08-990-379-4
; Sequence 4, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; TITLE OF INVENTION: Their Biologically Active Expression Products
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-08-990-379-4

```

Query Match	16.6%	Score	192.5	DB 2	Length	393			
Best Local Similarity	34.2%	Pred. No.	4.4e-14						
Matches	50	Conservative	29	Mismatches	58	Indels	9	Gaps	4
QY	57	EYWPVKLYIGDEATALDRYRLQAGETHVLNAAHGRNVDPGDPYRYRMDIQYHGVEADDL	116						
Db	197	EILPFLYLGSAHYAARMDLMDALGITALLNVS-----SDCPNHF-EGHYQYKCIPEVDN	249						
QY	117	PIFDLSVFYPPAAAFTRDALSDHDSKILYHCVMGRSRSATLVLAYLMIHKDMTLVDIAIQ	176						
Db	250	HRADISWFMEATEYID-AVKDCGRVLVHCQAGISRATCICLAYLMKKRVRLEAEAF	308						
QY	177	VAKNRCVL-PNRGFLKOLRELDKQV	201						
Db	309	VKQRRIISPNFSFMGQLQFESQVL	334						
RESULT	11								



Db 174 -EILPYLYGSCNHSDDLQGLQACGITTAVLNVS-----ASCPNHFEGL-PHYKSIPVED 225  
Qy 116 LPTDLSVEFYPAAFIDRALSDHSLVHCVMGSRSATLVLAYLMIHKDMLVDAIQ 175  
Db 226 NQWEISAWFOEATSFID-SVKNSGRVLVHCQAGISRSATLCIAYLIQSHRVRLEAFD 284  
Qy 176 QVAKNRCVL-PNRGFLKQLRELDKQLV 201  
Db 285 FVKORRGVISPNSFMGOLLQLETVL 311

## RESULT 14

US-09-164-193-22  
; Sequence 22, Application US/09164193C  
; Patent No. 6258582  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan L.  
; TITLE OF INVENTION: NOVEL CSATTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)  
; FILE REFERENCE: MNI-051  
; CURRENT APPLICATION NUMBER: US/09/164,193C  
; CURRENT FILING DATE: 1998-09-30  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-164-193-22

Query Match 15.9%; Score 184.5; DB 4; Length 314;  
Best Local Similarity 28.3%; Pred. No. 2.7e-13;  
Matches 56; Conservative 39; Mismatches 84; Indels 19; Gaps 6;

Qy 10 LKNAYSSAKRLSPKMEEGEEDYCTPGAFELER-----LFWKGSQYTHVNEVWPKLYI 64  
Db 128 LRGGFKSPQTCYCPDLCSEAPQAALPPAGAENSNDPRPIYDQGPV-----EILPYLYL 182  
Qy 65 GDEATLDYRLQKAGTHVLNAAHGRWNVDTPDYRDMDIQYHGVEADDLPFDLSVF 124  
Db 183 GCSNHSDDLQGLQACGITTAVLNVS-----ASCPNHFEGL-PHYKSIPVEDNQWVEISAW 235  
Qy 125 FYPAAFIDRALSDHSLVHCVMGSRSATLVLAYLMIHKDMLVDAIQVAKNRCVL 184  
Db 236 FQEAISFID-SVKNSGRVLVHCQAGISRSATLCIAYLIQSHRVRLEAFDFVKQRRGVI 294  
Qy 185 -PNRGFLKQLRELDKQLV 201  
Db 295 SPNFSFMGOLLQLETVL 312

## RESULT 15

US-09-045-973-5  
; Sequence 5, Application US/09045973  
; Patent No. 6165767  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah  
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA: US/09/045,973  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0491 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRSTNOT16  
CLONE: 3041794  
US-09-045-973-5

Query Match 15.3%; Score 177.5; DB 4; Length 198;  
Best Local Similarity 28.5%; Pred. No. 8.9e-13;  
Matches 53; Conservative 33; Mismatches 61; Indels 39; Gaps 6;

Qy 20 LSPKMEEGEEDYCTPGAFELERLFWKGSQYTHVNEVWPKLYIGDEATLDYRLQKA 79  
Db 14 MAPRMISEGD-----IGGIAQITSSSLFLGRGSVANRHLQAR 51  
Qy 80 GFTHVLNAA-----HGRWNVDTPDYRDMDIQYHGVEADDLPFDLSVFFYPAAAFIDRA 135  
Db 52 GITCIVNATIEIPNFN-----POF-----EYKVPPLADMPHAPIGLYFDTVADKI-HS 99  
Qy 136 LSDSHSLVHCVMGSRSATLVLAYLMIHKDMLVDAIQVAKNRCVL-PNRGFLKQLR 194  
Db 100 VSRKHGATLVHCAAGVSRSATLCIAYLMKPHNVCLLEAYNNVKARRPVIRENVGVFWRQLI 159  
Qy 195 ELDKQL 200  
Db 160 DYERQL 165

Search completed: February 8, 2002, 15:42:02  
Job time: 121 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:43:46 ; Search time 23.54 Seconds  
(without alignments)  
692.274 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 220

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VOQRRSQRDGEEDGREL 220

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size: 10

Total number of hits satisfying chosen parameters: 11

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	61.4	220	22	Human SGP003 phosph
2	98	44.5	221	21	A human regulator
3	12	5.5	57	22	Human SGP014 phosph
4	12	5.5	80	22	Human polypeptide
5	12	5.5	207	22	Human hydrolyase-li
6	12	5.5	211	21	Human dual specifi
7	12	5.5	211	22	Human polypeptide
8	12	5.5	211	22	Human SGP060 phosph
9	12	5.5	263	21	Human cardiovascular
10	12	5.5	328	22	Human polypeptide
11	12	5.5	549	22	Human SGP014 phosph

#### ALIGNMENTS

RESULT 1  
AAE04837  
ID AAE04837 standard; Protein: 220 AA.  
XX  
AC AAE04837;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human SGP003 phosphatase polypeptide.  
XX  
KW Human: SGP003 phosphatase polypeptide; phosphatase-related disease;  
KW immune-related disorder; ocular disease; organ transplant rejection;  
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;  
KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;  
KW neuroprotective; antibacterial; vulnery; tranquiliser; antiasthmatic;  
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;  
KW antifungal; dual specificity phosphatase; BSP; MAP kinase phosphatase;  
KW MKP; migraine; chromosome CHR10.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..173  
FT /label= Catalytic\_domain  
FT Domain 54..199  
FT /label= Phosphatase\_domain  
FT Misc-difference 85  
FT /note= "Encoded by CPG"  
XX  
PN WO200146394-A2.  
XX  
PD 28-JUN-2001.  
XX  
PF 21-DEC-2000; 2000WO-US34736.  
XX  
PR 21-DEC-1999; 99US-0173255.  
PR 28-DEC-1999; 99US-0175766.  
PR 25-JAN-2000; 2000US-0178078.  
PR 31-JAN-2000; 2000US-0179301.  
XX  
PA (SUGE-) SUGEN INC.  
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
PI Flanagan P;  
XX  
DR WPI; 2001-418058/44.  
DR N-PSDB; AAD09495.  
XX  
PT Novel phosphatase polypeptide useful for treating cancers,  
PT immune-related diseases and disorders, cardiovascular disease, brain or  
PT neuronal-associated diseases and metabolic disorders  
XX  
PS Claim 7; Fig 2; 186pp; English.  
XX  
CC The present invention relates to phosphatase polypeptides, nucleotide  
CC sequences encoding them, as well as various products and methods useful  
CC for the diagnosis and treatment of various phosphatase-related diseases  
CC and conditions. Substance that modulates the activity of phosphatase  
CC polypeptide is used to treat immune-related diseases and disorders,  
CC cardiovascular disease, brain or neuronal-associated diseases and  
CC metabolic disorders, including cancers of tissues, cancers of  
CC haematopoietic origin, diseases of central and peripheral nervous  
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,  
CC amyotrophic lateral sclerosis, viral infections, infections caused by

CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, hypotension, hypertension, psychotic disorders, neurological  
 CC amino acid sequence is human SGP003 phosphatase polypeptide. This  
 CC sequence is classified as dual specificity phosphatase (DSP) and MAP  
 CC kinase phosphatase (MKP). SGP003 gene maps to chromosomal position  
 CC CHR10.  
 XX  
 SQ Sequence 220 AA;

Query Match 61.4%; Score 135; DB 22; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-128;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 86 NAAHGRNWNVDGPDYRMDIQYHGVEADDLPTFDSLVFFPAAAFIDRALSDHRSKILV 145  
 DB 86 naahgrwnvdtgpyrdmdiqyhqveaddlptfalsvffpaaafidralsdhskilv 145  
 QY 146 HCVMGSRSSATLVLAYLMIHKDMTLVDAIQVAKNRCVLPNRGFLKQRLDQLVQQR 205  
 DB 146 hcvmgrrsatslvlaylmihkdmtilvdaiqvaknrcvlpnrgflkqireldkqlvqqr 205  
 QY 206 RSQRDGEEDGREL 220  
 DB 206 rsqrqgeedgrel 220

RESULT 2  
 AAB18667  
 ID AAB18667 standard; Protein; 221 AA.  
 AC AAB18667;

22-JAN-2001 (first entry)

A human regulator of intracellular phosphorylation.

KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;  
 KW neurological disorder; Parkinson's disease; demyelinating disease;  
 KW meningitis; developmental disorder; neuromuscular disorder; cancer;  
 KW myasthenia gravis; cell proliferative disorder; actinic keratosis;  
 KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;  
 KW autoimmune disorder; inflammatory disorder; Addison's disease;  
 KW acquired immunodeficiency disease; allergy; diabetes mellitus;  
 KW rheumatoid arthritis; microbial infection; trauma.

OS Homo sapiens.

Key Location/Qualifiers  
 FT Modified-site 2 /note= "potential phosphorylation site"  
 FT Modified-site 9 /note= "potential phosphorylation site"  
 FT Modified-site 16 /note= "potential phosphorylation site"  
 FT Modified-site 21 /note= "potential phosphorylation site"  
 FT Modified-site 69 /note= "potential phosphorylation site"  
 FT Modified-site 170 /note= "potential phosphorylation site"  
 FT Modified-site 208 /note= "potential phosphorylation site"  
 FT Active-site 146..158 /note= "tyrosine specific protein phosphatase active site"

W0200055332-A2.

21-SEP-2000.

PF 17-MAR-2000; 2000WO-US07277.  
 XX 18-MAR-1999; 99US-0125593.  
 PR 20-MAY-1999; 99US-0135049.  
 PR 09-JUL-1999; 99US-0143188.  
 PA (INCY-) INCYTE PHARM INC.  
 XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
 PI Lu DAN, Au-Young J;  
 XX WPI: 2000-602121/57.  
 DR N-PSDB: AAA75684.  
 XX Novel intracellular phosphorylation regulator polypeptides and  
 FT polynucleotides for diagnosis, prevention and treatment of  
 FT neurological, cell proliferative and autoimmune/inflammatory disorders  
 FT .  
 XX Claim 1; Page 84-85; 96pp; English.  
 CC The present sequence represents a human regulator of intracellular  
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
 CC are useful for treating a disease or condition associated with  
 CC decreased or increased expression of functional HRIP. Diseases treated  
 CC or diagnosed include neurological disorders such as stroke, Parkinson's  
 CC disease, demyelinating diseases, bacterial and viral meningitis and  
 CC other developmental disorders of the central nervous system.  
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
 CC inflammatory disorder such as Addison's disease, acquired  
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
 CC rheumatoid arthritis, microbial infection and trauma.  
 XX  
 SQ Sequence 221 AA;

Query Match 44.5%; Score 98; DB 21; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-90;  
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MTSGEVTSLKNAYSKRLSPKMEEGEEDYCTGAFELERLFWGSPQYTHVNEWVP 60  
 DB 1 mtsgevtslknayssakrlspkmeegeedyctgafelerlfwgspqythvnewvp 60  
 QY 61 KLYIGDEATLDRLYRLQKAGFTHVLNAAHGRWNVDGTP 98  
 DB 61 klyigdeatldryrlqkagfthvlnaahgrwnvdtgtp 98

RESULT 3  
 AAE04849  
 ID AAE04849 standard; Protein; 57 AA.  
 AC AAE04849;  
 XX 10-SEP-2001 (first entry)  
 DE Human SGP014 phosphatase polypeptide related exon 3.  
 KW Human; SGP014 phosphatase polypeptide; phosphatase-related disease;  
 KW immune-related disorder; ocular disease; organ transplant rejection;  
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
 KW attention disorder; cognition disorder; psychotic disorder; cyostatic;  
 KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;  
 KW neuroprotective; antibacterial; vulnerary; tranquiliser; antisthmatic;  
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;

KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
KW MKP; migraine; chromosome 10q21.3; exon 3.  
XX  
OS Homo sapiens.  
PN WO200146394-A2.  
XX  
PD 28-JUN-2001.  
XX  
PF 21-DEC-2000; 2000WO-US34736.  
XX  
PR 21-DEC-1999; 99US-0173255.  
PR 28-DEC-1999; 99US-0175766.  
PR 25-JAN-2000; 2000US-0178078.  
PR 31-JAN-2000; 2000US-0179301.  
XX  
PA (SUGE-) SUGEN INC.  
XX  
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
PI Flanagan P;  
XX  
DR WPI: 2001-418058/44.  
XX  
XX Novel phosphatase polypeptide useful for treating cancers,  
PT immune-related diseases and disorders, cardiovascular disease, brain or  
PT neuronal-associated diseases and metabolic disorders  
XX  
PS Example 1; Page 127; 186pp; English.  
XX  
CC The present invention relates to phosphatase polypeptides, nucleotide  
CC sequences encoding them, as well as various products and methods useful  
CC for the diagnosis and treatment of various phosphatase-related diseases  
CC and conditions. Substance that modulates the activity of phosphatase  
CC polypeptide is used to treat immune-related diseases and disorders,  
CC cardiovascular disease, brain or neuronal-associated diseases and  
CC metabolic disorders, including cancers of tissues, cancers of  
CC hematopoietic origin, diseases of central and peripheral nervous  
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,  
CC amyotrophic lateral sclerosis, viral infections, infections caused by  
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,  
CC sexual dysfunction, mood disorders, attention disorders, cognition  
CC disorders, hypotension, hypertension, psychotic disorders, neurological  
CC disorders, dyskinesias and organ transplant rejection. The present amino  
CC acid sequence is human SGP014 phosphatase polypeptide related exon.  
CC SGP014 sequence is classified as dual specificity phosphatase (DSP) and  
CC MAP kinase phosphatase (MKP). SGP014 gene maps to chromosomal position  
CC 10q21.3.  
XX  
SQ Sequence 57 AA;  
Query Match 5.5%; Score 12; DB 22; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 152 SRSATLVLAYLM 163  
Db 11 srsatlvlaylm 22  
|||||  
RESULT 4  
ID AAM41142  
XX AAM41142 standard; Protein; 80 AA.  
XX  
AC AAM41142;  
XX  
XX 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6073.  
XX  
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI60298.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries  
XX  
PS Example 2; SEQ ID NO 6073; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 80 AA;  
Query Match 5.5%; Score 12; DB 22; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.00022;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 152 SRSATLVLAYLM 163  
Db 26 srsatlvlaylm 37  
|||||  
RESULT 5  
ID AAB28793  
XX AAB28793 standard; Protein; 207 AA.  
XX  
AC AAB28793;  
XX  
XX 13-FEB-2001 (first entry)  
XX  
DE Human hydrolase-like molecule 4 protein.

XX Hydrolase-like molecule; human; cell proliferation disorder;  
KW autoimmune; cancer; AIDS; acquired immune deficiency syndrome.  
XX Homo sapiens.  
XX US6132964-A.  
PD 17-OCT-2000.  
XX 06-FEB-1998; 98US-0013881.  
XX 06-FEB-1998; 98US-0013881.  
XX (INCY-) INCYTE PHARM INC.  
PI Bandman O, Hillman JL, Guegler KJ, Shah P, Lal P, Corley NC;  
XX WPI; 2001-006133/01.  
DR N-PSDB; AAC60226.  
XX New human hydrolase-like molecules (HHLMs) and polynucleotides encoding  
PT the HHLMs, useful for diagnosing, treating or preventing cell  
PT proliferation (e.g. bone cancer) or autoimmune disorders(e.g. AIDS or  
PT asthma)  
XX Claim 1; Column 47-48; 38pp; English.  
PS The present invention relates to isolated and purified cDNA encoding a  
CC human hydrolase-like molecule (HHLM), designated HHLM-1 to HHLM-8. The  
CC HHLM DNAs and polypeptides are useful for diagnosing, treating or  
CC preventing cell proliferation disorders and autoimmune disorders. Cell  
CC proliferation disorders include cancers and autoimmune disorders include  
CC AIDS (acquired immune deficiency syndrome). The present sequence  
CC is a HHLM protein of the invention.  
XX Sequence 207 AA;

Query Match 5.5%; Score 12; DB 22; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163  
Db 153 srsatlvlaylm 164  
|||||

RESULT 6  
AAV85620  
ID AAV85620 standard; Protein; 211 AA.  
XX  
AC AAV85620;

XX 06-FEB-2001 (first entry)  
XX Human dual specificity phosphatase-9 (DSP-9) amino acid sequence.  
XX  
KW Dual specificity phosphatase-9; DSP-9; human; cytostatic; cell growth;  
KW immunosuppressive; antiallergic; antiproliferative; autoimmune disease;  
KW cancer; graft-versus-host disease; allergy; metabolic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200060100-A1.  
XX  
PD 12-OCT-2000.  
XX  
PF 07-APR-2000; 2000WO-US09321.  
XX  
PR 07-APR-1999; 99US-0128203.  
XX  
PA (CEPT-) CEPTVR INC.

XX Luche RM, Wei B;  
XX WPI; 2000-656232/63.  
DR N-PSDB; AAC61100.  
XX

Dual specificity phosphatase-9 which dephosphorylates activated  
mitogen-activated protein kinase, used to identify agents that inhibit  
DSP-9 activity and modulate cell proliferation, differentiation, and  
survival

Claim 1; Fig 2A; 66pp; English.

This invention relates to an isolated dual specificity phosphatase-9  
(DSP-9) and its variant. The DSP-9 protein has the ability to  
dephosphorylate an activated mitogen activated protein (MAP) kinase.  
Included in the invention are an expression vector comprising a  
polynucleotide encoding the DSP-9 protein, a host cell transformed by the  
expression vector, and an antibody that specifically binds to DSP-9.  
DSP-9 has cytostatic; immunosuppressive; antiallergic; and  
antiproliferative activity. DSP-9 modulating agents are useful for  
modulating a proliferative response, differentiation or survival of a  
cell which displays contacting inhibition of cell growth, anchorage  
independent growth or an altered intercellular adhesion property, in a  
patient. DSP-9 agonists and antagonists are also useful for treating a  
disorder associated with DSP-9 activity such as Duchenne muscular  
dystrophy, cancer, graft-versus-host disease, autoimmune disease,  
allergies, metabolic disease, abnormal cell growth, abnormal cell  
proliferation and cell cycle abnormalities. The present sequence  
CC represents the human DSP-9 protein.  
XX Sequence 211 AA;

Query Match 5.5%; Score 12; DB 21; Length 211;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163  
Db 157 srsatlvlaylm 168  
|||||

RESULT 7  
AAM39356  
ID AAM39356 standard; Protein; 211 AA.  
XX  
AC AAM39356;

XX 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 2501.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-052317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.  
 DR N-PSDB: AAI58512.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 4; SEQ ID NO 2501; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 211 AA;

Query Match 5.5%; Score 12; DB 22; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 0.00053;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 152 SRSATLVLAYLM 163  
 Db 157 srsatlvlaylm 168  
 |||||

RESULT 8  
 AAE04839  
 ID AAE04839 standard; Protein: 211 AA.

XX AAE04839;

XX 10-SEP-2001 (first entry)

XX Human SGP060 phosphatase polypeptide.

XX Human: SGP060 phosphatase polypeptide; phosphatase-related disease;  
 KW immune-related disorder; ocular disease; organ transplant rejection;  
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
 KW attention disorder; cognition disorder; psychotic disorder; cytostatic;  
 KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;  
 KW neuroprotective; antibacterial; vulnerrary; tranquilliser; antisthmatic;  
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;  
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
 KW MKP; migraine; chromosome 8p11.1-q11.1.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..173  
 FT /label= Catalytic\_domain  
 FT 61..204  
 FT /label= Phosphatase\_domain  
 XX

XX WO200146394-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US34736.

XX 21-DEC-1999; 99US-0173255.

XX 28-DEC-1999; 99US-0175766.

XX 25-JAN-2000; 2000US-0178078.

XX 31-JAN-2000; 2000US-0179301.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
 PI Flanagan P;

XX WPI: 2001-418058/44.

XX N-PSDB: AAD09497.

XX Novel phosphatase polypeptide useful for treating cancers,  
 PT immune-related diseases and disorders, cardiovascular disease, brain or  
 PT neuronal-associated diseases and metabolic disorders -

XX Claim 7; Fig 2; 186pp; English.

XX The present invention relates to phosphatase polypeptides, nucleotide  
 CC sequences encoding them, as well as various products and methods useful  
 CC for the diagnosis and treatment of various phosphatase-related diseases  
 CC and conditions. Substance that modulates the activity of phosphatase,  
 CC polypeptide is used to treat immune-related diseases and disorders,  
 CC cardiovascular disease, brain or neuronal-associated diseases and  
 CC metabolic disorders, including cancers of tissues, cancers of  
 CC haematopoietic origin, diseases of central and peripheral nervous  
 CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,  
 CC amyotrophic lateral sclerosis, viral infections, infections caused by  
 CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognitive  
 CC disorders, dyskinesias, hypertension, psychotic disorders, neurological  
 CC amino acid sequence is human SGP060 phosphatase polypeptide. This  
 CC sequence is classified as dual specificity phosphatase (DSP) and MAP  
 CC kinase phosphatase (MKP). SGP060 gene maps to chromosomal position  
 CC 8p11.1-q11.1.

XX Sequence 211 AA;

Query Match 5.5%; Score 12; DB 22; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 0.00053;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 152 SRSATLVLAYLM 163  
 Db 157 srsatlvlaylm 168  
 |||||

RESULT 9

AAAY92175

ID AAY92175 standard; Protein: 263 AA.

XX AAY92175;

XX 01-AUG-2000 (first entry)

XX Human cardiovascular system associated protein tyrosine phosphatase 2.

XX Cardiovascular system associated protein tyrosine phosphatase 2;  
 KW CSAPT-2; cytostatic; immunomodulatory; antidiabetic; virucide;

hypotensive; cardiant; tyrosine phosphatase modulator.  
Homo sapiens.  
Key Location/Qualifiers  
Domain 1..131 /label= N\_terminal\_domain  
/note= "unique"  
Modified-site 25..30 /note= "N-myristoylation site"  
Modified-site 11..13 /note= "Protein Kinase C phosphorylation site"  
Domain 31..43 /note= "helix-loop-helix DNA binding domain"  
Modified-site 47..53 /note= "tyrosine kinase phosphorylation site"  
Modified-site 50..53 /note= "N-glycosylation site"  
Modified-site 58..63 /note= "N-myristoylation site"  
Domain 76..263 /note= "phosphatase catalytic active domain"  
Modified-site 89..92 /note= "Casein kinase II phosphorylation site"  
Modified-site 112..117 /note= "N-myristoylation site"  
Modified-site 131..134 /note= "Casein kinase II phosphorylation site"  
Modified-site 160..165 /note= "N-myristoylation site"  
Modified-site 165..168 /note= "Casein kinase II phosphorylation site"  
Modified-site 176..179 /note= "amidation site"  
Modified-site 220..222 /note= "Protein Kinase C phosphorylation site"  
Modified-site 223..225 /note= "Protein Kinase C phosphorylation site"  
Modified-site 246..251 /note= "N-myristoylation site"  
Modified-site 257..260 /note= "Casein kinase II phosphorylation site"  
WO200018890-A2.  
06-APR-2000.  
30-SEP-1999; 99WO-US22924.  
30-SEP-1998; 98US-0163833.  
30-SEP-1998; 98US-0164193.  
(MILL-) MILLENNIUM PHARM INC.  
Acton S;  
WPI; 2000-293136/25.  
N-PSDB; AAA09027, AAA09028.  
New cardiovascular system associated protein tyrosine phosphatases are used in the diagnosis and treatment of e.g. immune disorders, anti-proliferative disorders, metabolic disorders and cardiovascular disorders  
Claim 8; Page 146-147; 156pp; English.  
The CSATPP nucleic acid and protein molecules are used to modulate regulation of cellular processes. CSATPP nucleic acid and protein molecules and modulators of CSATPP activity and expression can be used to treat a subject with a disorder characterized by aberrant CSATPP expression or activity. These disorders can include an immune disorder, an anti-proliferative disorder, a proliferative disorder e.g. renal and lung carcinomas, a metabolic disorder e.g. diabetes, viral pathogenesis,

a neural disorder, a cardiovascular disorder e.g. hypertension or coronary heart disease or a disorder arising from improper phosphorylation of a phosphorylated protein. Anti-CSATPP antibodies are used to isolate CSATPP by standard techniques, to facilitate the purification of natural and recombinantly produced CSATPP from cells, to detect CSATPP protein in cell supernatant for evaluating the level of CSATPP expression and to monitor protein levels in tissue as part of a clinical testing procedure to determine efficacy of a treatment procedure. CSATPP nucleic acids, antibodies and protein molecules can be used in screening assays and in predictive medicine e.g. prognostic assays, monitoring clinical trials and pharmacogenetics.  
xx Sequence 263 AA;  
SQ  
Query Match 5.5%; Score 12; DB 21; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 152 SRSATLVLAYLM 163  
|||||  
Db 114 srsatlvlaylm 125  
|||||  
RESULT 10  
AAM42355  
ID AAM42355 standard; Protein; 328 AA.  
XX  
AC AAM42355;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 88.  
XX  
KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery; antiparkinsonian; antischling; antianemic; antithratic; cancer;  
KW antirheumatic; hepatotropic; cerebrotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200155449-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01346.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUL-2000; 2000US-0216880.  
PR 14-JUL-2000; 2000US-0218290.  
PR 14-AUG-2000; 2000US-0225447.  
PR 01-SEP-2000; 2000US-0229343.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0231243.  
PR 25-SEP-2000; 2000US-0234997.  
PR 29-SEP-2000; 2000US-0236367.  
PR 13-OCT-2000; 2000US-0239937.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246528.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249265.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:45:01 ; Search time 15.43 Seconds  
(without alignments)  
1086.092 Million cell updates/sec

Title: US-09-847-519A-2  
Perfect score: 220  
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VOORRRSQRDGEEDGREL 220

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 10

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
---------------	-------	----------------	--------	-------	-------------

No matches found

Search completed: February 8, 2002, 15:48:10  
Job time: 189 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:46:07 ; Search time 11.72 seconds  
(without alignments)  
688.248 Million cell updates/sec

Title: US-09-847-519A-2  
Perfect score: 220  
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRQDGEEDGREL 220

Scoring table: OLIGO  
Gapop 60.0' , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 10

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----	-----	-----	-----	-----	-----

No matches found

Search completed: February 8, 2002, 15:49:20  
Job time: 193 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: February 8, 2002, 15:45:47 ; Search time 23.66 seconds  
(without alignments)  
1360.098 Million cell updates/sec

Title: US-09-847-519A-2  
Perfect score: 220  
Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRQGEEDGREL 220

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 10

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL\_17:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	5.5	86	4 Q9BTW0	Q9btw0 homo sapien
2	12	5.5	198	11 Q9D700	Q9d700 mus musculu
3	12	5.5	211	4 Q9BV47	Q9bv47 homo sapien

ALIGNMENTS

RESULT 1  
Q9BTW0  
ID Q9BTW0 PRELIMINARY; PRT; 86 AA.  
AC Q9BTW0;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE UNKNOWN (PROTEIN FOR MGC:2627).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=NEUROBLASTOMA;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003115; AA031115.1;  
SQ SEQUENCE 86 AA; 9652 MW; 4DD29AFD5989528B CRC64;

Query Match 5.5%; Score 12; DB 4; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLYLM 163  
|||||  
Db 32 SRSATLVLYLM 43

RESULT 2  
Q9D700  
ID Q9D700 PRELIMINARY; PRT; 198 AA.  
AC Q9D700;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE 2310043K02RIK PROTEIN.  
GN 2310043K02RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Kochiwa H.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kasaiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
"Functional annotation of a full-length mouse cDNA collection.";  
RT Nature 409:685-690(2001).  
DR EMBL; AK009781; BAB26501.1;  
DR MGD; MGI:1914209; 2310043K02RIK.  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR Pfam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.  
DR SMART; SM00012; PTPC\_DSPC; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolase.  
SQ SEQUENCE 198 AA; 22419 MW; 11F0CAF4B5620F0E CRC64;

Query Match 5.5%; Score 12; DB 11; Length 198;  
Best Local Similarity 100.0%; Pred. No. 0.00035;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163  
|||||  
Db 144 SRSATLVLAYLM 155

## RESULT 3

Q9BV47 PRELIMINARY; PRT; 211 AA.  
AC Q9BV47;  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE UNKNOWN (PROTEIN FOR MGC:1136).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG CARCINOMA;  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC001613; AAR01613.1; -;  
SQ SEQUENCE 211 AA; 23945 MW; 60E944304905086D CRC64;

Query Match 5.5%; Score 12; DB 4; Length 211;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SRSATLVLAYLM 163  
|||||  
Db 157 SRSATLVLAYLM 168

Search completed: February 8, 2002, 15:49:02  
Job time: 195 sec



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	199.4	17.1	594	4	US-09-163-833-3	Sequence 3, Appli
2	199.4	17.1	912	4	US-09-163-833-1	Sequence 1, Appli
3	175	15.0	1691	3	US-09-013-881-12	Sequence 12, Appl
4	161.6	13.9	789	4	US-09-164-133-6	Sequence 6, Appli
5	161.6	13.9	1016	4	US-09-164-133-4	Sequence 4, Appli
6	133	11.4	861	1	US-07-988-273-1	Sequence 1, Appli
7	133	11.4	861	5	PCT-US93-12019-1	Sequence 1, Appli
8	70.6	6.1	1238	2	US-08-530-230-11	Sequence 11, Appl
9	60.2	5.2	1987	2	US-08-990-379-1	Sequence 1, Appli
10	60.2	5.2	1993	2	US-08-990-379-2	Sequence 2, Appli
11	51.6	4.4	4451	3	US-08-717-294-42	Sequence 42, Appl
12	49.8	4.3	2097	3	US-08-941-445A-10	Sequence 10, Appl
13	48.6	4.2	1176	2	US-08-387-942C-17	Sequence 17, Appl
14	48.6	4.2	1258	2	US-08-387-942C-1	Sequence 1, Appli
15	48.2	4.1	11219	1	US-07-643-734C-1	Sequence 1, Appli
16	48.2	4.1	11219	3	US-08-439-009A-1	Sequence 1, Appli
17	47.6	4.1	1729	4	US-09-045-973-6	Sequence 6, Appli
18	47.2	4.1	7218	1	US-08-232-463-14	Sequence 14, Appl
19	47	4.0	1704	1	US-08-528-199-2	Sequence 2, Appli
20	47	4.0	1704	1	US-08-528-199-5	Sequence 5, Appli
21	46.2	4.0	5515	4	US-09-398-193-98	Sequence 98, Appl
22	46.2	4.0	30001	1	US-08-125-468-1	Sequence 1, Appli
23	46.2	4.0	30001	2	US-08-474-933-1	Sequence 1, Appli
24	45.8	3.9	1557	5	PCT-US91-01327-12	Sequence 12, Appl
25	45.8	3.9	1557	6	5248670-2	Patent No. 5248670
26	45.6	3.9	2943	1	US-08-042-747A-7	Sequence 7, Appli
27	44.8	3.8	2825	4	US-09-196-390-5	Sequence 5, Appli

RESULT 3  
US-09-013-881-12  
; Sequence 12, Application US/09013881  
; Patent No. 6132964  
; GENERAL INFORMATION:

Qy	231	gcctttgagctgagcggctctctctggaaggcagccccagtacaccacagctcaacagag	290
Db	544	gtcttcgagttggagcggctctctacacaggcaagacagcctgttaaccatgccgacgag	603
Qy	291	gtctggcccaagctctacattgctgcatggaagcagcggcgtggaacgctataggctgcag	350
Db	604	gtctggccaggcctctatctcgagaccaggacatggctaacaaacgccgggagacttcgc	663
Qy	351	aaggcgggggtctcgcacgtctgtgaacggccacggcccgctgggaactggaactggg	410
Db	664	cgcttgggcatacgcacgctctctcaatgcttcacacagccggtggcgag-----gcacg	717
Qy	411	cccgactactaccgcgacatggacatccagtaccacggcggtaggcgcgacgacctgcc	470
Db	718	cccgaggccctataggggctgggcataccgcgtacctgggtttga-----gccca	765
Qy	471	accttgacctgaagtctctctctaccggcggagccttcacacagagcgttaagc	530
Db	766	gcctttgacatgacatccacttccacagcggctggcgacttcattccacggggcgctgagc	825
Qy	531	gacgaccacagtaagatccctggttcactgctcatgggcgcgacgggtcagccacacctg	590

Db	826	CAGCCAGGAGGAAGATCCTGCTGTCATTGCTGTGGCGGTGAGCCGATCGCCACCCCTG	885
Qy	591	gtcctggcctacctgatgatccacaagacatgacccctgggtggcgcacatccagcaagtg	650
Db	886	GTACTGGCCTACTCATGCTGTACACACCACTTACCTTCGTGGAGGCCATCAAGAAATC	945
Qy	651	gccaaagaaccgtctgcgtctctcccaaccggggcctttttaagcagctcccgagctggac	710
Db	946	AAAGACCACGAGGATCATCCCAACCGGGGCTTCCTGAGGAGCTCCTGGCCCTGGAC	1005
Qy	711	aagcagctggtgcag	725
Db	1006	CGCAGGCTGCGCAG	1020

```

RESULT      4
US-09-164-193-6
; Sequence 6, Application US/09164193C
; Patent No. 6258582
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSATTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
; FILE REFERENCE: MNI-051
; CURRENT APPLICATION NUMBER: US/09/164,193C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
US-09-164-193-6

```

Query Match	13.9%;	Score 161.6;	DB 4;	Length 789;
Best Local Similarity	61.9%;	Pred. No. 4.5e-28;		
Matches 276;	Conservative 0;	Mismatches 164;	Indels 6;	Gaps
QY	245	gcggctctcttgaaaggcagtcgccaggtacaccacgctcaacagaggtctgcccacaaagt	304	
Db	18	gcggctcctctacacaggaacagcctgtaaccatgccacgaggtctgcccaggcct	77	
QY	305	ctacattggcgatgaggcgacgcgctgaccgcgtataggctgcagaaaggcgggttcac	364	
Db	78	ctatctcgagacagagacatggctaacaccgcgggagcttcgcgcctggggcatcac	137	
QY	365	gcacgtgctgaacgcggccacgcgcgttgaaacgtggacactggcccgactactacacg	424	
Db	138	gcacgtcctcaatgctctcacacagccggtggcgag-----gcacgcccgagcctatga	191	
QY	425	cgacatggacatccagtatacaaggcgtggaggccgcagacttgcacactctgcacctcag	484	
Db	192	ggggctgggcatcgcgtactactgggtgtgaggccacgactgcgcagcctttgacatgag	251	
QY	485	tgtctctctataccggcggcgcagccttcctatgcagagcgctaaagcgcagaccacagtaa	544	
Db	252	cattccacttcagagcgtgcgcacttcctaccacggcgctgagccagccagaggagaa	311	
QY	545	gattcttggttacctgcgtcatgggcgcgcagcggctcagccacctggtcttgacctacct	604	
Db	312	gattcttggtgcatgtgctgtgggcgtgagccgcatcgcgcaccttggtaactggcctacct	371	
QY	605	gatgtaccacaagacatgaccttggtagcgcctatccagcaagtggccaaagaaccctgg	664	
Db	372	catgtgtaccaccacttacctctgtggaggccatacaagaaagtcaagaccaccgag	431	
QY	665	cgctctcccgaaacggggctttttga	690	
Db	432	agagggccgagcccgagccactatca	457	

```

RESULT      5
US-09-164-193-4
; Sequence 4, Application US/09164193C
; Patent No. 6258582
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID
; FILE REFERENCE: MNI-051
; CURRENT APPLICATION NUMBER: US/09/164,193C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1016
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
; US-09-164-193-4

```

	Query Match	13.9%;	Score 161.6;	DB 4;	Length 1016;
	. Best Local Similarity	61.9%;	Pred. No. 4.8e-28;		
	. Matches	276;	Conservative	0; Mismatches 164;	Indels 6; Gaps
Qy	245	gcggctcttctggaaggccagtcgcccaagtacaccacgtcaacgcagaggtctgcccagaagt	304		
Db	18	gcggctctctacacagcgaagacagcgtctaaccatgccgacgaggtctggccaggcct	77		
Qy	305	ttacattggcgatgagcgacgcgcgtggaccctataggctgcagaaggcggggttcac	364		
Db	78	ctactcggagaccaggacatggtctaacaccgcgggagcttcgcgcgtggggcatcac	137		
Qy	365	gcacgtgtgaacggccacagccgctggaaactggacactggcccgacactactacog	424		
Db	138	gcacgtctcaatgcttcacacagccggtggcag-----gcacgcccgagccctatga	191		
Qy	425	cgcacatggacatcagtacacagcgtggaggccgcagcactgccacactctgcacctag	484		
Db	192	ggggtcgtgcatccgtactcgtgggtgttgaggccacgactcccgactcttgacatgag	251		
Qy	485	tgtctttctataccggcggcagcccttcactgcacagagcgtcaagcgacgaccacagtaa	544		
Db	252	catccactccagcgtccgacttcattccaccggcgctgagccagccaggaggaa	311		
Qy	545	gatccctggttcactgcctcatgggcgcgacccggtcagccacccctggtctcgccctacct	604		
Db	312	gatccctggtcatttgtctgtgggcgtgagccgcatcccgccacccctggtaetggccctacct	371		
Qy	605	gatgatccacaagacatgacctggtgagcgccatccacgaagtggccagagaccgcgtg	664		
Db	372	catgctgtaccaccacactaccctcgtggaggccatacaagaagtcaaaagaccaccgagg	431		
Qy	665	cgctctccgaacccggggtcttttga	690		
Db	432	agaagccgagccccagccactgttca	457		

RESULT 6  
US-07-988-273-1  
; Sequence 1, Application US/07988273  
; Patent No. 5512434  
; GENERAL INFORMATION:  
; APPLICANT: AARONSON, Stuart A.  
; APPLICANT: BOTTARO, Donald P.  
; APPLICANT: ISHIBASHI, Toshio  
; APPLICANT: MIKI, Toru  
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN  
; TITLE OF INVENTION: PHOSPHATASE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:



APPLICANT: Hughes, David Anthony  
TITLE OF INVENTION: Methods for Screening of Substances for  
TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/530,290  
FILING DATE: 14-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB94/00694  
FILING DATE: 31-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9402573.1  
FILING DATE: 10-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9307250.2  
FILING DATE: 07-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 084611-00000005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1238 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-530-290-11

Query Match 6.1%; Score 70.6; DB 2; Length 1238;  
Best Local Similarity 52.9%; Pred. No. 2.3e-07;  
Matches 176; Conservative 0; Mismatches 154; Indels 3; Gaps 1;  
Qy 393 tgaacgtgacactggccgactactaccgacatgacatccagaccagcggtg 452  
Db 729 TTGAATGTCCTCGGACTGCCAACCCTTTGAAGGACACTATCAGTACAACTGCATC 788  
Qy 453 gagccgacgacactggccacacttgcacactcagtgctcttctaccgcgcccttc 512  
Db 789 CCACTGGAAGATACCAACCAAGCGGCACATCAGCTCCTGTTTCATGGAAGCATAGATAC 848  
Qy 513 atcgacagagcgttaagcagaccacagtaagatcctggttactcgtcgtatggccgc 572  
Db 849 ATCGA---TGCCGTTGAAGGACTGCGGTGGCGCGTGTGTGTCACATGCCAGCGGGCATC 905  
Qy 573 agccggtcagccacctggcttggcctactatgatccacagacatgacacctggtg 632  
Db 906 TCGCGGTGCGCCACCATCTGCTGGCCTACCTGATGATGAGAAACGGGTGAGCTGGAG 965  
Qy 633 gacgccatccagcaagtggccaaacacgctgctcctccggaaccggggctttttgaag 692  
Db 966 GAGGCCCTTCGAGTTGCTTGAAGCAGCGCGCAGCATCATCTCGCCCACTTCAGCTTCATG 1025  
Qy 693 cagctcgggagctggaacagcagctggtgcag 725  
Db 1026 GGCAGCTGTGTCAGTTTCGAGTCCAGGTGCTG 1058

## RESULT 9

US-08-990-379-1  
Sequence 1, Application US/08990379  
Patent No. 5998188  
GENERAL INFORMATION:  
APPLICANT: Stork, Philip J  
APPLICANT: Misra-Press, Anita  
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and  
TITLE OF INVENTION: Their Biologically Active Expression Products  
FILE REFERENCE: 4104-000322USA  
CURRENT APPLICATION NUMBER: US/08/990,379  
CURRENT FILING DATE: 1997-12-15  
EARLIER APPLICATION NUMBER: PCT/US96/10402  
EARLIER FILING DATE: 1996-06-14  
EARLIER APPLICATION NUMBER: 60/000,263  
EARLIER FILING DATE: 1995-06-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 1987  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-08-990-379-1

Query Match 5.2%; Score 60.2; DB 2; Length 1987;  
Best Local Similarity 51.1%; Pred. No. 6.2e-05;  
Matches 168; Conservative 0; Mismatches 158; Indels 3; Gaps 1;  
Qy 395 gaacgtggacactggccgactactaccgacatgacatccagaccagcggtgga 454  
Db 1052 gaatgtctctcagactgcccccaatcactttgaggacattaccagtcacagtcacccc 1111  
Qy 455 ggcgcgacgacctgccacacttgcacactcagtgcttcttctaccgcgccgacacctcat 514  
Db 1112 ggtagaagataaccacaaagctgacatcagctcctggttcattggaagccatcgaatacat 1171  
Qy 515 cgacagagcgttaagcagacaccacacagtaagatcctggttactcgtcgtcagcgacag 574  
Db 1172 agac---gcagtgaaggactgcccaggcgagtgctggttactcagtcgacggtcctc 1228  
Qy 575 ccggtcagccacctggtcctggcctacctgatgacacagacatgacacgacacctggtgga 634  
Db 1229 tagatcagccacctgctcctggcctacctgatgacacagacggtgaggtcgtgga 1288  
Qy 635 cgccatccagcaagtggccaaacacgctgctcctcccgaaaccggggctttttgaagca 694  
Db 1289 ggcttgcagttcgtcaagcagcgccgtagcatcatctcgcaccaacttcagcttcattggg 1348  
Qy 695 gctcgggagctggaacagcagctggtgc 723  
Db 1349 ccagttgctgcagttcgcagttcaggtgc 1377

## RESULT 10

US-08-990-379-2  
Sequence 2, Application US/08990379  
Patent No. 5998188  
GENERAL INFORMATION:  
APPLICANT: Stork, Philip J  
APPLICANT: Misra-Press, Anita  
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and  
TITLE OF INVENTION: Their Biologically Active Expression Products  
FILE REFERENCE: 4104-000322USA  
CURRENT APPLICATION NUMBER: US/08/990,379  
CURRENT FILING DATE: 1997-12-15  
EARLIER APPLICATION NUMBER: PCT/US96/10402  
EARLIER FILING DATE: 1996-06-14  
EARLIER APPLICATION NUMBER: 60/000,263  
EARLIER FILING DATE: 1995-06-16  
NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1993  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-08-990-379-2

Query Match 5.2%; Score 60.2; DB 2; Length 1993;  
Best Local Similarity 51.1%; Pred. No. 6.2e-05;  
Matches 168; Conservative 0; Mismatches 158; Indels 3; Gaps 1;  
QY 395 gaacgtgagacactggccgactactaccgacatgacatccagaccacgagcggtgga 454  
Db 1058 gaatgtctctcagactgcccactactctttgaggacattaccagatcaagtcatccc 1117  
QY 455 ggcgagacactgcccactctgacatcagtgcttcttcttaccgagcgagccttcat 514  
Db 1118 ggtagaagataaacacacaggtgacatcagctcctggttcatggaagccatcgatacat 1177  
QY 515 cgacagagcgctaaagcgacacacagataagatcctgttctacgtcgtcatggtggccgag 574  
Db 1178 agac--gcagtagagactgcccagagggagtgctgttctactgcccagccggcatctc 1234  
QY 575 ccggtcagcaccctgctcctggcctaccctgatgacacacagacatgacccctggtgga 634  
Db 1235 tagatcagccaccatctgctggcctaccctgatgacacacagagagagcggtgaggtggagga 1294  
QY 635 cgccatccagcaagtgcccaagacccgctgctcctcccaagcgggtttttgagaca 694  
Db 1295 ggcttctgagtcgaagcagcgctgacatcatctcgcacactcagcttcatggtg 1354  
QY 695 gctccggagactgacacagcagctggtgc 723  
Db 1355 ccagttgctcagttcagttcaggtgc 1383

## RESULT 11

US-08-717-294-42  
Sequence 42, Application US/08717294  
Patent No. 6114148  
GENERAL INFORMATION:  
APPLICANT: SEED, BRIAN  
APPLICANT: HAAS, JURGEN  
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/717,294  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/345001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

TELEX:  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4451 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-717-294-42

Query Match 4.4%; Score 51.6; DB 3; Length 4451;  
Best Local Similarity 44.9%; Pred. No. 0.0072;  
Matches 195; Conservative 0; Mismatches 239; Indels 0; Gaps 0;  
QY 328 cgctggaccgctagctgcagaagcggggttcacgacgtgctgaacgagcgccacacg 387  
Db 34 CGCTGGCCACCTTGTACTGCTGGGGATGCTGGTCTCGTTCGCTAGCCGCCACCCGCC 93  
QY 388 gcccgtggaacgtggacactggccgacactactaccgacatgacatccagaccacg 447  
Db 94 GCTACTACCTGGGCGCGTGGAGCTGCTCTGGGACTACATCAGAGCGACCTGGCGGAGC 153  
QY 448 gctggagggcgacgacactgcccacacttcgacctcagtgctcttcttaccgcgggcag 507  
Db 154 TCCCGGTGGAGCGCGCTTCCCGCCCGCGTGGCCAAAGAGCTTCCCTTCAACACACGCG 213  
QY 508 ccttcacgcagagcgctaaagcagacacacagatcctggttctacgtcgtcatgg 567  
Db 214 TGGTGTAAAGAAACCTGTTCTGGAGTTACCGCACCTGTTCAACATTGCCAAGC 273  
QY 568 gcccagccggtcagccacactgctcctggcctaccctgatgacacacagacatgaccc 627  
Db 274 CGCGCCCGCTGATGGCGCTGCTGGCGCCACCATCCAGCGCGAGGTGTACACACCG 333  
QY 628 tggtagcgccatccacgacagtgcccaagaaacccgctgctcctccgaaacgggctttt 687  
Db 334 TGGTGATCACCTGAAGAATGCGCCACCGCCCGTCCAGCTGCAGCGCTGGCGGTGA 393  
QY 688 tgaagcgctccggagcgctggacaaagcagctggtgacgacagcgagcggtccacagcc 747  
Db 394 GCTACTGGAGGCGCAGCGCGCGCGAGTACGACGACGACGACGACGACGACGACG 747  
QY 748 aggcagcgtaggag 761  
Db 454 AGGACGACAGGTG 467

## RESULT 12

US-08-941-445A-10  
Sequence 10, Application US/08941445A  
Patent No. 6107060  
GENERAL INFORMATION:  
APPLICANT: Keeling, Peter  
APPLICANT: Guan, Hanping  
TITLE OF INVENTION: Starch Encapsulation  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/941,445A  
FILING DATE: 30-SEP-1997  
CLASSIFICATION: 800

```

RESULT 14
US-08-387-942C-1
Sequence 1, Application US/08387942C
Patent No. 5939289
GENERAL INFORMATION:
APPLICANT: ERTESVAG, HELGA
APPLICANT: VALLA, SVEIN
APPLICANT: SKJAK-BRAEK, GUDMUND
APPLICANT: LARSEN, BOORN
TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: P.O BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/387,942C  
APPLICATION NUMBER: 09-MAY-1995  
FILING DATE: 09-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY JR, GERALD M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1809-106P  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12588 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE: Azotobacter vinelandii  
STRAIN: E  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 290..1951  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2227..6438  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6702..9695  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 9973..12588  
US-08-387-942C-1

Query Match 4.28; Score 48.6; DB 2; Length 12588;  
Best Local Similarity 47.8; Pred. No. 0.048;  
Matches 141; Conservative 0; Mismatches 154; Indels 0; Gaps 0;  
Qy 375 aacgcggccacgcccgtggaacgtgacactggcccaactactactacccgacatgac 434  
Db 782 AGCTGGCCCGACGACGAGCGGCTTCGACGGCTTCGTCGCGGACTACCTGGTGCACGCGTG 841  
Qy 435 atccagtaccacgcccgtgagcgacgacctgccaccttcgacctcagtgctttctc 494  
Db 842 TTCGAGAACACGCTCGCTACGCCAACGACCGCCACGCTTCACAGTGGTCACGACGACC 901  
Qy 495 taccgcggcggcagccttcacgcagagcggtacgacgacaccacagtaagatcctggtt 554  
Db 902 CACGATTTCGTCATGACCAACAGCGTCCCTACGCGCAACGCGACGCGGCGCTGGTGTG 961  
Qy 555 cactgcgtcatggcgcgacccggtcagccacctggttcctggcctaccctgatgatccac 614  
Db 962 CAGCGGGGTCTGAGGAGGACCTCGCGCTGCCAGCAACATCCTGTGACGCGGCGGCGCTAC 1021  
Qy 615 aaggacatgacccctggtggacgacctcagcaagtggcccaagaacccgctgcgtcc 669  
Db 1022 TACGACACGCGCCGGAAGGCGTCTGCTCAAGATGACGACGACATCACCTCGC 1076

RESULT 15  
US-07-642-734C-1  
Sequence 1, Application US/07642734C  
Patent No. 5824513  
GENERAL INFORMATION:  
APPLICANT: Katz, L  
APPLICANT: Donadio, S  
APPLICANT: McAlpine, J B  
TITLE OF INVENTION: Recombinant DNA Method for Producing

TITLE OF INVENTION: Erythromycin Analogs  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edward H. Gorman  
STREET: Abbott Laboratories D377/AP6D-2 One Abbott  
CITY: Park Rd  
STATE: IL  
COUNTRY: US  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/642,734C  
FILING DATE: 17-JAN-91  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dankers, Andreas M  
REGISTRATION NUMBER: 32652  
REFERENCE/DOCKET NUMBER: 4952 US.01  
TELEPHONE: 708-937-9396  
TELEFAX: 708-938-2623  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11219 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Saccharopolyspora erythraea  
STRAIN: NRRL 2338  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 744..6659  
OTHER INFORMATION: /function= "APPROXIMATE SPAN OF  
OTHER INFORMATION: MODULE 1"  
OTHER INFORMATION: /label= FUNCTION  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 744..11219  
OTHER INFORMATION: /function= "gene= "eryA"  
OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for  
OTHER INFORMATION: 5-deoxyerythronolide B"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 744..1868  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: acyltransferase domain 1 of module 1"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1998..2198  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: acyl carrier domain 1 of module 1"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2250..3626  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3831..4811  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: acyltransferase 2 domain of module 1"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 5574..6125



```
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6369..6826
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain 2 of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6678..11219
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6678..8066
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacyl ACPsynthase of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8262..9305
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9906..10454
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10707..10964
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 2"
;
; US-07-642-734C-1
```

```
Query Match 4.1%; Score 48.2; DB 1; Length 11219;
Best Local Similarity 45.9%; Pred. No. 0.057;
Matches 164; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 362 cagcagcgtgtaagcagcgccgacgctggaacgtggaacactggcgccgactacta 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9950 CGCGAGGTGGCCCGGTGGCTCCCGGTGCGCGCGCGGACACCTCGCGTGGTCAGCAG 10009

QY 422 ccgagacatgacatccagtacacagcggtggagggcgacgacctgccaccttcgacct 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10010 GCGAGGCCGACACCGAGGGCGTCCGCGACCTGACCGCCGAGCTGACCGCGCTCGGCGC 10069

QY 482 cagtgctttcttaccggcgagccttcctcgcagagagcgctaaagcagaccacag 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10070 GCGGGTGTGCGGTGCACGCGTGCAGCGTACAGCGCGGACCGGAGGAACTCGTGCA 10129

QY 542 taagatcctggttcactgcgtatggtggcgagcggtgagccacctggtcctgccta 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10130 CGGCCCTGATCGACAGGGCGGCGTGCCTCCCGGTGGTGCACGCGGCGGACTGCGCA 10189

QY 602 cctgatgatccaaagacatgacctggtggagcgccatccagcagtggtgccaagaccg 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10190 GCAGGTCCGATCACACATGACGACGAGCGCGCTTCGACGAGGTGGTCCGCGCCAAAGC 10249

QY 662 ctgctctctcccgaacggggtttttgaagcagctccgggagctggacaagcagct 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10250 CGGGGGCGCGGTGCACCTGGACGAGCTGTGCTCGGAGCGCGAGCTGTTCCTGCTGTT 10306
```

Search completed: February 8, 2002, 18:43:16  
Job time: 4241 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 18:14:05 ; Search time 1166.24 Seconds  
(without alignments)  
10734.353 Million cell updates/sec

Title: US-09-847-519A-1  
Perfect score: 1165  
Sequence: 1 ggcagtggtgggtggtggg.....aaaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estfun.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estom.\*  
5: em\_estpl.\*  
6: em\_estba.\*  
7: em\_estro.\*  
8: em\_estov.\*  
9: em\_htc.\*  
10: gb\_est1.\*  
11: gb\_est2.\*  
12: gb\_htc.\*  
13: gb\_gss.\*  
14: em\_gss\_fun.\*  
15: em\_gss\_hum.\*  
16: em\_gss\_inv.\*  
17: em\_gss\_pin.\*  
18: em\_gss\_pro.\*  
19: em\_gss\_rod.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.4	38.1	572	11	BF074326
2	425.8	36.5	634	10	AW918423
3	401	34.4	605	10	AW918455
4	368	31.6	478	11	BF652341
5	242.4	20.8	505	11	BE808287
6	200	17.2	486	11	BF430044
7	194.4	16.7	703	11	BF199835
8	190	16.3	1748	12	AK009781
9	187.6	16.1	558	11	BF077053
10	176.6	15.2	990	11	BF207232
11	166.2	14.3	1104	12	AK007513
12	165.8	14.2	471	11	BF774179

13	164	14.1	415	10	AI372800
14	163.6	14.0	1065	12	AK006247
15	161.8	13.9	456	10	BE751020
16	160.4	13.8	612	11	BI197674
17	159.8	13.7	469	10	AW425509
18	158.2	13.6	961	11	BF314818
19	153.6	13.2	693	11	BG083406
20	148.8	12.8	369	11	BF430184
21	148.4	12.7	607	10	AL585353
22	144.2	12.4	411	11	BF826456
23	143.6	12.3	641	11	BI393955
24	135.6	11.6	642	11	BI393954
25	134.4	11.5	1149	11	BI083188
26	133	11.4	550	11	BE892660
27	133	11.4	644	11	BG17129
28	133	11.4	723	10	BE387475
29	133	11.4	850	11	BE750047
30	133	11.4	949	10	BE383984
31	133	11.4	1001	10	AL555009
32	131.6	11.3	561	10	AL602266
33	129.6	11.1	461	11	BF776211
34	129.4	11.1	566	10	AI526580
35	128.8	11.1	946	11	BG246446
36	128.8	11.1	1169	12	AK008734
37	127.2	10.9	718	10	BE390181
38	125.8	10.8	545	10	BE757502
39	121.4	10.4	957	11	BF527844
40	118	10.1	713	11	BF212271
41	116.4	10.0	781	10	AU168945
42	114.8	9.9	609	10	AW951243
43	114.8	9.9	684	10	BE613888
44	113.2	9.7	430	10	BE771213
45	111.4	9.6	927	10	BE728819

ALIGNMENTS

RESULT 1  
BF074326 LOCUS BF074326 572 bp mRNA EST 25-APR-2001  
DEFINITION 221711 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION BF074326  
VERSION BF074326.1 GI:10867837  
KEYWORDS EST.  
SOURCE COW.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 572)  
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Portea,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.  
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle  
JOURNAL Genome Res. 11 (4), 626-630 (2001)  
MEDLINE 21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCAGTCACGACG  
Plate: 80 row: I column: 23







REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Bovidae; Bovinae; Bos.  
1 (bases 1 to 486)  
Warren, W.C., Teo, N., Allison, T., Wagner, S., Mathialagan, N., Kata, S.,  
Johnson, J., Smith, T.P.L. and Womack, J.E.  
A survey of genes transcribed in bovine skeletal muscle  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel.: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR primers  
FORWARD: GGAACAGCTATGACCATG  
BACKWARD: GTTTCCTCCAGTCACGAC  
Seq primer: AATTAACTCTCACTAAAGG.  
Location/Qualifiers  
1. .486  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC BSM"  
/tissue\_type="Skeletal muscle"  
/lab\_host="XL0LR"  
/note="Vector: Uni-ZAP XR; Site\_1: EcoRI; Site\_2: XhoI;  
Library obtained from Stratagene, catalog #937721. Library  
made from skeletal muscle of a two year old Holstein cow."  
96 a 147 c 136 g 107 t

## FEATURES

source

BASE COUNT

ORIGIN

Query Match 17.2%; Score 200; DB 11; Length 486;

Best Local Similarity 64.4%; Pred. No. 2.4e-26;

Matches 299; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

Qy 266 tccccagtagcaccacgctgagaggtgtgcccgaagctctacattgagtaggagcgac 325

Db 21 TGCCATGTTGAACCATTTCAATGAAGTCTGCCCAACCTCTTCTGGAGACGCTACGC 80

Qy 326 gggcgtgagcgtctataggtgcagaagggggttcacgacgctgtgaacgagcgccca 385

Db 81 AGCCCGGACAAGAACAAAGTTGACCCAGCTGGGCATCACACATGTTGTAATGTTGCCGC 140

Qy 386 cggcgtggaacgtgacactggcccgactactacggagacatcgagacatccagtagca 445

Db 141 AGGCAAGTTTCAGGTGGACACAGGTGCGCAAGTTCTACCGGGAATGCCCTTGGAGTACTA 200

Qy 446 cggcgtggaacgtgacactggccaccttgcacctcagtgctgtcttctacccgagcg 505

Db 201 TGGCATCGAGGCTGATGACAAACCCCTTCTTTGACCTCAGTGTCTACTTTCGCCCGTGC 260

Qy 506 agccttcacgacagagcgttaagcagcagcagcagtaagatccttggttcaactcgtcat 565

Db 261 TCGATATACGGAAGTGCCTCAGTGTGCCAAGCGCGGTGTGTACACTGTGCCAT 320

Qy 566 gggccgagcagcgttcaagccaccttggcttgacctactgatgacacagagacatgac 625

Db 321 GGGGGTGAAGCGCTCTGCCACAGTGTCTGGCCCTTCTCATGATCTCGGAGAACATGAC 380

Qy 626 cctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 685

Db 381 GCTGGTGAAGCCATCCAGACAGTGCAGGCCACCGCGATATCTGCCCCCAACTCGGGCTT 440

Qy 686 ttgaagcagcgtcgggagctggagcagcagcagcagcagcagcagcagcagcagcagc 729

Db 441 CCTCCGAGCTCCAGGCTCTGGACACACCGACTGGGGGGGAGA 484

RESULT 7

BI199835

LOCUS

BI199835 703 bp mRNA EST

10-JUL-2001

## DEFINITION

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

602761093F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4896479 5',  
mRNA sequence.

BI199835

BI199835.1 GI:14654856

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 703)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLC1783 row: 1 column: 24

High quality sequence stop: 701.

Location/Qualifiers

1. .703

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="IMAGE:4896479"

/clone\_lib="NIH\_MGC\_19"

/tissue\_type="neuroblastoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCAGGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-CDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT 139 a 223 c 223 g 118 t

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Length

703;

Qy 231 gctttgagctggagcgcctcttctggaaggcagtcctccagtagcacccacgtaacag 290

Db 38 GTCCTTCAGTTGGAGCGCTCTCTACACAGGCAAGACAGCGCTTAACCATGCCGACGAG 97

Qy 291 gctgcccgaagctctacattggcgatgagcgagcgctggaccgctataggtcgag 350

Db 98 GTCTGGCCAGGCGCTCTATCTCGGAGACGACATGGCTAACAAACCGCGGAGCTTCCC 157

Qy 351 aaggcggggttcacgacgtgtgaacgcccacgcccgtggaacgtggaacatggg 410

Db 158 GCCTTGGGCAATCAGCAGCTCTCAATGCCCTCACACAGCGGTGGCGAG-----GCACG 210

Qy 411 cccgactactaccgcgcacatggacatccagtaaccagcgctggagcgacgacctgcc 470

Db 211 CCCGAGCCCTATGAGGGCTGGGCATCCGCTACCTGGGTGTGAGGCCACGACTCCCA 270

Qy 471 accttcacctcagtgcttcttcttcttcttcttcttcttcttcttcttcttcttctt 530

Db 271 GCTTTGACATGAGCATCCACTTCCAGAGCGCTGCCGACTTCATCCACCGGCGCTGAGC 330

Qy 531 gacgaccacagtaagatccctgttcttcttcttcttcttcttcttcttcttcttctt 590

Db 331 CAGCCAGGAGGGAAGATCCTTGGTGTCTGTGGCGGTGAGCCCATCCGCCACCTGT 390

Qy 591 gctcgtacctcctgtagtcacaaaggaacatgaccttggtagcgccatccagcaagt 650











Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Ferlie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

**TITLE**  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

**JOURNAL**  
Nature 377 (8547 Suppl), 3-174 (1995)

**MEDLINE**  
96026280

**COMMENT**  
Other ESTs: EST175329 EST175330 EST175331 EST175332

EST175333 EST175334 EST175335 EST175336 EST175337 EST175338

EST175339 EST17534

Contact: HGI (Human Gene Index)

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-0200

Fax: (301)-838-0208

Email: hgi@tigr.org

**FEATURES**  
Location/Qualifiers

1..415

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="F0H1B1"

/lab\_host="E. coli DH5-alpha"

/note="Vector: BA, M13-derived; Site\_1: HindIII; Site\_2:

NotI; The infant brain library, constructed by Bento

Soares, Columbia University, was oligo-(dT) primed and

directionally cloned into an M13-derived plasmid using

total brain mRNA from a 72-day old human female afflicted

with spinal muscular atrophy."

**BASE COUNT** 82 a 132 c 123 g 71 t 7 others

**ORIGIN**

Query Match 14.1%; Score 164; DB 10; Length 415;

Best Local Similarity 63.3%; Pred. No. 7.1e-20;

Matches 264; Conservative 0; Mismatches 147; Indels 6; Gaps 1;

Qy 276 accacgtcaacagtgcttgcggcccaagctctacattggcgtgacgcgcgtggac 335

Db 5 AACCATCGGACGAGTCTGGCCAGGCTCTATCTCGGACACGACATGNTAACAC 64

Qy 336 cgcctataggctgcagaagcgggggttcacgcacgtgctgaaocggccacgcgcgtgg 395

Db 65 CCGCGGGAGCTTCGCGCGCTGGGCGATCACGCGCTCTCAATGCTCACAGCGCGTGG 124

Qy 396 aacgtggacactggccgcactactaccgcacatccagatccacacacgcgcgtggag 455

Db 125 NCAGGCAC-----GCCGAGGCCCTATGAGGGGNTGGGCATCCGCTACTGGGTGGAG 178

Qy 456 gccgcagcactgcccacacttcagctcagtgctctcttcacccggcgagccttcac 515

Db 179 GCCCAGGACTCGCCAGCCTTTGACATGAGCATCCACTTCACAGCGGTCCGACTTCATC 238

Qy 516 gacagagcgttaagcagacacacagtaagaatcctggttoactgcgtcatgcccgcagc 575

Db 239 CACCGGGCGCTGAGCGACGACGAGGAGAGATCCTGCTGCTGCTGCTGCTGCTGCTG 298

Qy 576 cggctcagcacccttgctcgtcctacactgatgatccacagagcatgaacctggtggac 635

Db 299 CGATCCGCCACCTGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358

Qy 636 gccatccagcagtgcccaagaacgcgtgcgtcctctcccgacccggggtttttgaag 692

Db 359 GGCATCAAGAAGTCAAGAACGACGAGGNATCATCCCAACCGGGGTCTCTGAGG 415

**RESULT** 14  
AK006247 1065 bp mRNA HTC 05-JUL-2001  
AK006247  
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700022L10, full insert sequence.

**DEFINITION**

**ACCESSION**

**VERSION**

**KEYWORDS**

**SOURCE**

Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:1700022L10.

**ORGANISM**

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**REFERENCE**

**AUTHORS**

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGATCCAGACTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGAGCGCGCAATTAATTCGAGTTAATTAATTAATCCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

## FEATURES

source location/Qualifiers  
1. .1065  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="MGI:1351599"  
/db\_xref="MGI:1901550"  
/clone="1700022L10"  
/sex="male"  
/tissue\_type="testis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
221. .817  
/note="putative"  
/codon\_start=1  
/protein\_id="BAB24480.1"  
/db\_xref="GI:12839241"  
HINFEWNLFLGDAYAARDGLRIOLGITHVYVNAAGKFOYDQAKRYGTPLLEYGI  
EADNPFFDLSVHPILPVARYIDALNIPRSRLVHCARGVSRSAITVLAFLMIFENMT  
LVDAIQTVQAHRDICPNSGFLRQLQVLNRRRETGR"

## BASE COUNT

245 a 301 c 290 g 229 t

## ORIGIN

Query Match 14.0%; Score 163.6; DB 12; Length 1065;  
Best Local Similarity 57.7%; Pred. No. 5.8e-20;  
Matches 292; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 231 gctttgagctgagcggctcttctggaaggcagtcacccagtcacccagtcacag 290  
|||||  
Db 305 GCCTCTCTGCGACGATGCTGTGGCTCGGCTCGGACTGCGACACTGACCCACATCAATGAG 364  
|||||  
QY 291 gctgccccaaagctctacatgctgagcggcagcggcgtcagcgtctagcgtcag 350  
|||||  
Db 365 GTCTGCGCCACACTTTCTTGGAGATGCGTATGCTGCGAGACAAAGGTCGTCTAATC 424  
|||||  
QY 351 aaggcggggttcacgcagctgctgaacgcggccacgcggcgtggaacgtgacactggg 410  
|||||  
Db 425 CAGCTGGGCATTACCCATGTTCTGAATGCTGGCTGCGGCGCAAGTTCAGGTGGACAGGT 484  
|||||  
QY 411 cccgactactaccgcagcatgacatgacatcagaccgcggcgtggaagccagacactgccc 470  
|||||  
Db 485 GCCAAGTTCTACCGTGGAAACACTCTGGAGTACTATGGCATTTGAGGCTGATGACCAACCC 544  
|||||  
QY 471 accttcgactcagtgcttctctaccgcggcggcagcgtctcagcagcagcgtcagc 530  
|||||  
Db 545 TTCTTTGACCTCAGCGTCCACTTTCTGCTGTTGCTGTATGCTGCTGCTGCTGCTCAAT 604  
|||||  
QY 531 gacgaccacagtaagatcctggctcactcgtcgtcatggcgcgcgcgcgcgcgcgcgc 590  
|||||  
Db 605 ATTCGCCGAAGCGAGTGTGCTGCTACCTGCGCTAGGGGGTGGTGGCTGCTGCTGCTCAAT 664  
|||||  
QY 591 gctctggcctactatgatcatcacaagacatgacccctgggagcgcgcgcgcgcgcgcgc 650  
|||||  
Db 665 GTCTTGGCCTTCTCATGATCTTCAGAAACATGACACTGGTATGATGCTGCTGCTGCTGCTG 724  
|||||  
QY 651 gccaaagaacccgctgctctcccccgaacgggcttttgaagcagcgtccggagctgac 710  
|||||  
Db 725 CAGGCCACCGAGATATCTGTCCCAACTCAGGCTTCTCCGACAGCTCCAGCTTCTGGAC 784  
|||||  
QY 711 aagcagctggtgacagagggcagc 736

Db 785 AACAGCTGAGCGCGGAACAGGAAG 810  
|||||

## RESULT 15

BE751020 456 bp mRNA EST 25-APR-2001  
LOCUS 202721 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.  
DEFINITION BE751020  
ACCESSION BE751020  
VERSION BE751020.1 GI:10165012  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
Bovidae; Bovinae; Bos.

## REFERENCE

1 (bases 1 to 456)  
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett  
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.,  
Perte, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and  
Keele, J.W.

## TITLE

Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle

## JOURNAL

Genome Res. 11 (4), 626-630 (2001)

## MEDLINE

21180013

## COMMENT

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.98904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

## PCR Primers

FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCCAGTCACGACG  
Plate: 42 row: P column: 11  
Seq primer: ATTTAGGTGACACTATAG.

## FEATURES

source location/Qualifiers  
1. .456  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"

/note="vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."

## BASE COUNT

81 a 148 c 135 g 92 t

## ORIGIN

Query Match 13.9%; Score 161.8; DB 10; Length 456;  
Best Local Similarity 62.7%; Pred. No. 1.7e-19;  
Matches 269; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 201 gggaggaggaggaactgactgacccctggagccttggagctggagcgtctcttcttggag 260  
|||||  
Db 29 GGAGATCGGAAGCCAGCCGCTTGGCCAGCGTCTTGGAACTGGAGGAGCTCTGTAGGGCA 88  
|||||  
QY 261 ggcagtcacccagtcacccagtcacccagtcacccagtcacccagtcacccagtcacccag 320  
|||||  
Db 89 GGAAGGTTCTTCCAGCCACGCTGGATGAAGTTTGGCCCACTTTACATAGAGATGCG 148  
|||||  
QY 321 ggcagcggcgtgacgcgtatagcgtcagaaagcggggttcacacagcgtgctgacgcg 380  
|||||  
Db 149 GCCAGCAAAATTAACCGCTTTGAGCTATGGAAGCTGGGCAATTAACCCAGCTGCTGATGCC 208  
|||||  
QY 381 gcccacggcgcgtggaacgtggacactggcccgactactaccgcgacatggacatccag 440  
|||||

```

Db 209 GCCCAGGGGGCTCTACTGTGAGGGCAGCCCTGACTTCTATGGCAGCAGTGTGAGC--- 265
QY 441 taccacggcgtgagcgcagacaccccttcgacctcagtgctcttcttaccgc 500
Db 266 TACCTGGGGGTCCAGCCACGACCTTCTGAGTTCGACATCAGTGTCTACTTCTCCTCT 325
QY 501 gggcagccttcacgcagagcgcgtgagcgcagacccacagtaagatcctggtcactgc 560
Db 326 GCAGCTGACTTATCCACCGTGCCCTTAGCAGCCCTGGGGCCCAAGGTCTCTGGTGCCTGC 385
QY 561 gtcatgggcccagcgcgtcagccacctgtctctgacctgatgatccacaaggac 620
Db 386 GTGGTCGGGGTGAGCCGCTCCGCCACACTGGTCTCTGGCTACCTCATGCTGCGCCAGCAG 445
QY 621 atgaccctg 629
Db 446 CTCtCCCTG 454

```

Search completed: February 8, 2002, 19:05:04  
Job time: 3059 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2002, 18:13:10 ; Search time 104.24 Seconds  
(without alignments)  
9581.580 Million cell updates/sec

Title: US-09-847-519A-1

Perfect score: 1165

Sequence: 1 ggcacgtggggggcgctggg.....aaaaaaaaaaaaaaaaaaaa 1165

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 42862619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N\_Geneseq\_1101.\*
- 1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.\*
  - 2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.\*
  - 3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.\*
  - 4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.\*
  - 5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.\*
  - 6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.\*
  - 7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.\*
  - 8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.\*
  - 9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.\*
  - 10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.\*
  - 11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.\*
  - 12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.\*
  - 13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.\*
  - 14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.\*
  - 15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.\*
  - 16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.\*
  - 17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.\*
  - 18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.\*
  - 19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.\*
  - 20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.\*
  - 21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.\*
  - 22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1096.6	94.1	1262	22	Human SGP003 phosph
2	1071	91.9	1212	21	DNA encoding a hum
C 3	241	20.7	556	22	Probe #9870 used t
C 4	234	20.1	234	22	Probe #22910 used
5	204.2	17.5	636	22	Human SGP060 phosph
6	204.2	17.5	1878	22	Human polynucleoti
C 7	203.4	17.5	1691	22	Human cDNA SEQ ID
8	202.6	17.4	1491	21	Human cDNA encodin
9	201	17.3	1917	22	Human SGP014 phosph
10	199.4	17.1	904	22	Human phosphatase
11	199.4	17.1	1200	21	Human ORFX ORF2098

12	197.8	17.0	663	21	AAA09032
13	197.8	17.0	928	21	AAA09031
14	197.8	17.0	1023	21	AAZ46148
15	197.8	17.0	1300	21	AAC2749
16	196.2	16.8	1380	22	AAI59133
17	175	15.0	1691	22	AAC60226
18	165.2	14.2	597	22	AAF63566
19	161.6	13.9	789	21	AAA09028
20	161.6	13.9	1016	21	AAA09027
21	145.8	12.5	4360	22	AAO9494
22	133	11.4	861	15	AAQ68661
23	133	11.4	4580	21	AAC77082
24	123	10.6	451	21	AAC75128
25	119.4	10.2	240	22	AAO9503
26	110	9.4	110	21	AAC16446
27	105.6	9.1	775	21	AAC61101
28	105.6	9.1	1556	22	AAI60298
29	103.4	8.9	7733	22	AAI62831
30	98	8.4	1227	22	AAI59134
31	96.4	8.3	812	22	AAI60919
32	96.4	8.3	812	22	AAI60920
33	70.6	6.1	1238	15	AAQ72864
34	70.6	6.1	1238	16	AAQ6010
35	70.6	6.1	1238	22	AAF68858
36	70.6	6.1	2064	22	AAF68877
37	70.6	6.1	2109	22	AAF68878
38	69	5.9	1619	22	AAF68856
39	69	5.9	4637	22	AAF68859
C 40	68.4	5.9	4580	21	AAC77082
C 41	66.4	5.7	476	22	AAI15696
C 42	66.4	5.7	476	22	AAI37368
C 43	65.2	5.6	230	22	AAI24920
C 44	65.2	5.6	230	22	AAI50550
45	63.2	5.4	35828	21	AAA29063

ALIGNMENTS

RESULT 1

AAD09495  
ID AAD09495 standard; DNA; 1262 BP.

XX AAD09495;

XX 10-SEP-2001 (first entry)

XX Human SGP003 phosphatase polypeptide encoding DNA.

XX Human: SGP003 phosphatase polypeptide; phosphatase-related disease; immune-related disorder; ocular disease; organ transplant rejection; infection; diabetes; pain; sexual dysfunction; Alzheimer's disease; metabolic disorder; haematopoietic cancer; mood disorder; cardiac; Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; cardiovascular disease; brain; neuronal-associated disease; dyskinesia; attention disorder; cognition disorder; psychotic disorder; cytostatic; neurological disorder; virucide; nootropic; cerebroprotective; therapy; neuroprotective; antibacterial; vulnerrary; tranquiliser; antisthmatic; hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive; antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase; MKP; migraine; Chromosome CHR10; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 240..902

XX /\*tag= a

XX /product= "Human SGP003 phosphatase polypeptide"

XX /transl\_except= (pos:492..494, aa:leu)

XX WO200146394-A2.

XX 28-JUN-2001.





rhumatoid arthritis; microbial infection; trauma; ss.  
Homo sapiens.  
Key Location/Qualifiers  
CDS 172..837  
/\*tag= a  
/product= "regulator of intracellular phosphorylation"  
WO200005332-A2.  
21-SEP-2000.  
17-MAR-2000; 2000WO-US07277.  
18-MAR-1999; 99US-0125593.  
20-MAY-1999; 99US-0135049.  
09-JUL-1999; 99US-0143188.  
XX (INCY-) INCYTE PHARM INC.  
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
PI Lu DAM, Au-Young J;  
XX WPI; 2000-602121/57.  
DR P-PSDB; AAB18667.  
XX Novel human intracellular phosphorylation regulator polypeptides and  
PT polynucleotides for diagnosis, prevention and treatment of  
PT neurological, cell proliferative and autoimmune/inflammatory disorders  
PT  
PT  
PT  
XX Claim 4; Page 94; 96pp; English.  
XX The present sequence encodes a human regulator of intracellular  
CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
CC are useful for treating a disease or condition associated with  
CC decreased or increased expression of functional HRIP. Diseases treated  
CC or diagnosed include neurological disorders such as stroke, Parkinson's  
CC disease, demyelinating diseases, bacterial and viral meningitis and  
CC other developmental disorders of the central nervous system.  
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
CC including leukemia, melanoma, myeloma and cancer of the adrenal gland,  
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
CC inflammatory disorder such as Addison's disease, acquired  
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
CC rheumatoid arthritis, microbial infection and trauma.  
XX  
XX Sequence 1212 BP; 284 A; 345 C; 373 G; 210 T; 0 other;  
Query Match 91.9%; Score 1071; DB 21; Length 1212;  
Best Local Similarity 97.4%; Pred. No. 4.4e-230;  
Matches 1131; Conservative 0; Mismatches 25; Indels 5; Gaps 4;  
Qy 2 gccagtgagggtgctggcgctggcgctgtacatgccccacgacacgaacacctcccgac 61  
Db 54 gccagtgagggtgctggcgctggcgctgtacatgccccacgacacgaacacctcccgac 113  
Qy 62 gcggccagggcccgccacacccagctgcagaaaggagagaaatcccttgctctaaaat 121  
Db 114 gcggccagggcccgccacacccagctgcagaaaggagagaaatcccttgctctaaaat 173  
Qy 122 gacatctggagaagtgagacaaacctcaagaatgcttactatctgccaagaggtgtc 181  
Db 174 gacatctggagaagtgagacaaacctcaagaatgcttactatctgccaagaggtgtc 233  
Qy 182 gccagaatgag 241  
Db 234 gccagaatgag 293

## RESULT 3

AAI41184/C

ID AAI41184 standard; DNA; 556 BP.

XX AAI41184;

AC AAI41184;

XX 17-OCT-2001 (first entry)

XX

XX

XX











[illegible]



CC erythematous, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
SQ Sequence 1200 BP; 251 A; 364 C; 335 G; 248 T; 2 other;

Query Match 17.1%; Score 199.4; DB 21; Length 1200;  
Best Local Similarity 64.3%; Pred. No. 2.1e-35;  
Matches 299; Conservative 0; Mismatches 166; Indels 0; Gaps 0;  
QY 276 accacgctcaacagaggttgcgccaagctctacattgctgcatgagcgacagcgctgac 335  
DB 490 aaccatctcgatgaggttgcgcccgcctctctcctggagatgctacgacgcccggac 549  
QY 336 cgtatagctgcagaaaggcggttcacgcacgtgctgaacgcgccacgcccgtgg 395  
DB 550 aagagcaagctgacagctgggaatcacccacgcttgatgcgctgaggaagttc 609  
QY 396 aacgtggacactggccgactactacgcgcacatggacatccagtagcagcggtgag 455  
DB 610 caggtggacacaggtgcgaattctaccgtggaatgctccctggagtagctatggcatcgag 669  
QY 456 gccagacacctggccacacctgacactgaatgtcttcttaccggcgagccttcate 515  
DB 670 ggggagacacaccccttctgacactcagtgctacttctgctgttgcgtacacac 729  
QY 516 gacagagcgtcaagcagacacacagtagatcctggttctactgctgctacatggcgagc 575  
DB 730 cgagctgcccctcagtggttcccccaaggcgcgtgctggtacactgtgcatggggaagc 789  
QY 576 cgttcagcaccctggtctgctgactacatgacatgcacacaggaacatgacactggtagc 635  
DB 790 cgtctgcacactgtctgctgctctctcatgactatgatagaacatgacgctgtagag 849  
QY 636 gccatccagcagtggtgccaagaacgctgctcctccgaaacgggggttttgaagcag 695  
DB 850 gccatccagcagtggtgagccacacgcaatctctgcccactacactcaggttctctccgagc 909  
QY 696 ctccggggagctgagcaacagcagctggtgagcagagggcagcggtcc 740  
DB 910 ctccagggttctggacaacacgactggggcgagagcggggcggttc 954

## RESULT 12

AAA09032  
ID AAA09032 standard; DNA; 663 BP.  
XX  
AC AAA09032;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
DE Human CSATP-4 open reading frame.

XX Cardiovascular system associated protein tyrosine phosphatase 4;  
KW CSATP-4; cytosolic; immunomodulatory; antidiabetic; virucide;  
KW hypotensive; cardiant; tyrosine phosphatase modulator; ss.

XX Homo sapiens.

XX WO200018890-A2.

XX 06-APR-2000.

XX 30-SEP-1999; 99WO-US29294.

XX 30-SEP-1998; 98US-0163833.

XX 30-SEP-1998; 98US-0164193.

XX (MILL-) MILLENNIUM PHARM INC.

XX

PI Acton S;  
XX  
DR WPI; 2000-293136/25.  
DR P-PSDB; AAY92177.

PT New cardiovascular system associated protein tyrosine phosphatases are  
PT used in the diagnosis and treatment of e.g. immune disorders,  
PT anti-proliferative disorders, metabolic disorders and cardiovascular  
PT disorders

XX Claim 1; Page 155-156; 156pp; English.

XX This sequence encodes a cardiovascular system associated protein tyrosine  
CC phosphatase 4 (CSATP). The CSATP nucleic acid and protein molecules  
CC are used to modulate regulation of cellular processes. CSATP nucleic  
CC acid and protein molecules and modulators of CSATP activity and  
CC expression can be used to treat a subject with a disorder characterized  
CC by aberrant CSATP expression or activity. These disorders can include  
CC an immune disorder, an anti-proliferative disorder, a proliferative  
CC disorder e.g. renal and lung carcinomas, a metabolic disorder e.g.  
CC diabetes, viral pathogenesis, a neural disorder, a cardiovascular  
CC disorder e.g. hypertension or coronary heart disease or a disorder arising  
CC from improper phosphorylation of a phosphorylated protein. Anti-CSATP  
CC antibodies are used to isolate CSATP by standard techniques, to  
CC facilitate the purification of natural and recombinantly produced CSATP  
CC from cells, to detect CSATP protein in cell supernatant for evaluating  
CC the level of CSATP expression and to monitor protein levels in tissue as  
CC part of a clinical testing procedure to determine efficacy of a treatment  
CC procedure. CSATP nucleic acids, antibodies and protein molecules can be  
CC used in screening assays and in predictive medicine e.g. prognostic  
CC assays, monitoring clinical trials and pharmacogenetics.

XX Sequence 663 BP; 126 A; 222 C; 182 G; 133 T; 0 other;

Query Match 17.0%; Score 197.8; DB 21; Length 663;  
Best Local Similarity 64.1%; Pred. No. 4e-35;  
Matches 298; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 276 accacgctcaacagaggttgcgccaagctctacattgctgcatgagcgacgctgac 335  
DB 199 aaccatctcgatgaggttgcgcccgcctctctcctggagatgctacgacgcccggac 258  
QY 336 cgtatagctgcagaaaggcggttcacgcacgtgctgaacgcgccacgcccgtgg 395  
DB 259 aagagcaagctgacagctgggaatcacccacgcttgatgcgctgaggaagttc 318  
QY 396 aactggacactggccgactactacgcgcacatggacatccagtagcagcgctgag 455  
DB 319 caggtggacacaggtgcgaattctaccgtggaatgctccctggagtagctatggcatgag 378  
QY 456 gccagacacctggccacaccttcgacactcagtgcttcttctaccggcgagccttcate 515  
DB 379 gcggacgacaacccctctctcgacactcagtgcttctgctgttgcgtacatc 438  
QY 516 gacagagcgttaagcgcagcacacagtaagatcctggttctactgctcatggcgagc 575  
DB 439 cgagctgcctcagtggttcccccaaggcgctgctgtgtacactgtgcatggggttaagc 498  
QY 576 cgttcagcaccctggttctgctgactatgatccacaaggaacatgacacctggtgagc 635  
DB 499 cgtctgcacactgtcctgcttctcctcatgatcatgatgataacatgacgtgtagag 558  
QY 636 gccatccagcagtggtgcgaagaacacgctgctcctccgaacccgggttttgaagcag 695  
DB 559 gccatccagcagtggtgcagggccacccaataatctgccttaactcaggttctctccggcag 618  
QY 696 ctccggggagctgggacaacagcagctggtgcagcagagggcgagcggtcc 740  
DB 619 ctccaggttctggacaacacgactggggcgagcagggggcggttc 663

## RESULT 13







Search completed: February 8, 2002, 18:45:16  
Job time: 1926 sec



GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:41:46 : Search time 23.44 Seconds  
(without alignments)  
1372.863 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSAKRL.....VOORRRSRQDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_17.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	39.7	198	11 Q9QYJ7	Q9QYJ7 mus musculus
2	460.5	39.7	198	4 Q9UII6	Q9UII6 homo sapien
3	455	39.2	198	11 Q9DA25	Q9DA25 mus musculus
4	445	38.3	211	4 Q9BV47	Q9BV47 homo sapien
5	440	37.9	198	11 Q9D700	Q9D700 mus musculus
6	345.5	29.8	245	5 Q9VWN2	Q9VWN2 drosophila
7	335	28.9	185	11 Q9DTX3	Q9DTX3 mus musculus
8	233	20.1	86	4 Q9BTW0	Q9BTW0 homo sapien
9	205	17.7	184	4 Q9NRW4	Q9NRW4 homo sapien
10	204.5	17.6	365	5 Q41128	Q41128 caenorhabdi
11	203.5	17.5	198	10 Q9ER37	Q9ER37 arabisopsis
12	203.5	17.5	198	10 Q9LUG6	Q9LUG6 arabisopsis
13	201.5	17.4	946	10 Q9FFA8	Q9FFA8 arabisopsis
14	200	17.2	184	11 Q9GN11	Q9GN11 mus musculus
15	199.5	17.2	167	10 Q9MBK7	Q9MBK7 arabisopsis
16	193	16.6	436	11 Q9KCK2	Q9KCK2 mus musculus
17	193	16.6	1045	5 Q9NKL1	Q9NKL1 drosophila
18	192.5	16.6	303	4 Q9NSW1	Q9NSW1 homo sapien
19	192.5	16.6	394	4 Q13524	Q13524 homo sapien

20	192.5	16.6	411	4 Q13649	Q13649 homo sapien
21	190	16.4	661	10 Q9ATY4	Q9ATY4 zea mays (m
22	185	15.9	784	10 Q9C5S1	Q9C5S1 arabisopsis
23	182	15.7	608	5 Q9VU80	Q9VU80 drosophila
24	181.5	15.6	369	13 Q91790	Q91790 xenopus lae
25	179	15.4	665	4 Q9BY84	Q9BY84 homo sapien
26	179	15.3	690	4 Q9C0G3	Q9C0G3 homo sapien
27	177.5	15.3	198	4 Q95147	Q95147 homo sapien
28	176.5	15.2	353	13 Q42253	Q42253 gallus gall
29	174.5	15.0	482	4 Q9Y6W6	Q9Y6W6 homo sapien
30	174	15.0	220	11 Q9D6P6	Q9D6P6 mus musculus
31	174	15.0	220	11 Q99N12	Q99N12 mus musculus
32	174	15.0	223	11 Q60970	Q60970 mus musculus
33	174	15.0	223	11 Q9DCF8	Q9DCF8 mus musculus
34	173.5	14.9	476	5 Q9VHV8	Q9VHV8 drosophila
35	173.5	14.9	483	11 Q9CZY9	Q9CZY9 mus musculus
36	173	14.9	677	11 Q99MG6	Q99MG6 mus musculus
37	172.5	14.9	177	11 Q9CSL5	Q9CSL5 mus musculus
38	172.5	14.9	198	11 Q9D715	Q9D715 mus musculus
39	172.5	14.9	483	11 Q9ESS0	Q9ESS0 mus musculus
40	170.5	14.7	476	5 Q46122	Q46122 drosophila
41	169.5	14.6	189	11 Q9D9D8	Q9D9D8 mus musculus
42	169.5	14.6	198	11 Q9JLY7	Q9JLY7 mus musculus
43	169	14.6	348	5 Q9VVM4	Q9VVM4 drosophila
44	168	14.5	205	11 Q60969	Q60969 mus musculus
45	168	14.5	226	5 Q93592	Q93592 caenorhabdi

ALIGNMENTS

RESULT 1

Q9QYJ7

ID Q9QYJ7 PRELIMINARY; PRT; 198 AA.

AC Q9QYJ7; 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE PROTEIN PHOSPHATASE (DUAL-SPECIFICITY PHOSPHATASE TS-DSP6).

GN DUSP13.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20053896; PubMed=10585869;

RA Nakamura K., Shima H., Watanabe M., Haneji T., Kikuchi K.;

RT "Molecular cloning and characterization of a novel dual-specificity protein phosphatase possibly involved in spermatogenesis.";

RL Biochem. J. 344:819-825(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Aoyama K., Matsuda T., Aoki N.;

RT "Molecular cloning of a novel dual specificity phosphatase.";

KL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF027003; BAA89411.1; -

DR EMBL: AB237620; AAK15037.1; -

DR HSSP: P51452; 1VHR.

DR MGI:1351599; Dusp13.

DR InterPro: IPR000340; DS\_phosphatase.

DR InterPro: IPR000387; TYR\_phosphatase.

DR Pfam: PF00782; DSPC; 1.

DR SMART: SM00195; DSPC; 1.

DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.

DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.

DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.

KW Hydrolase.

SQ SEQUENCE 198 AA; 22481 MW; C2493597D6D3732B CRC64;

Query Match 39.7%; Score 461; DB 11; Length 198;  
Best Local Similarity 49.7%; Pred. No. 5.4e-36;



```

ID Q9BV47 PRELIMINARY; PRT; 211 AA.
AC Q9BV47;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MCC:1136).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC001613; AAH01613.1;
SQ SEQUENCE 211 AA; 23945 MW; 60E944304905086D CRC64;

Query Match 38.3%; Score 445; DB 4; Length 211;
Best Local Similarity 54.3%; Pred. No. 1.9e-34;
Matches 89; Conservative 26; Mismatches 47; Indels 2; Gaps 1;

Qy 39 FELERLFWKSPQYTHVNEVMPKLYIGDEATLDRYLQKAGFTHLVNAAGHGMVNDTGP 98
| | | | | : | : | | | | | : | : | | | | | : | : | | | | |
Db 46 FELERLLYTGKTACHNADEVMPGLYLGQDMANNRRRLRGITHTLVNASHNRW--GTP 103
| | | | | : | : | | | | | : | : | | | | | : | : | | | | |
Qy 99 DYIRDMIOYHGVEADLPTFDLSVFFYPAAAFIDRALSDHSHKILVHCVMGRSRSATLV 158
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Db 104 EAYEGIGIRYLGVEAHDSPAFDMSIHFTQAAFDIHRALSQPGGKTLVHCACVGSRSATLV 163
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Qy 159 LAYLMIHKDMTLVDAIQVAKNRCVLPNRFGLKQLRELDKQLVQ 202
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Db 164 LAYLMLYHHLTLVEAIKKVKDHGRIIPNRFGLQLLALDRRLRQ 207
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |

RESULT 5
Q9D700 PRELIMINARY; PRT; 198 AA.
ID Q9D700
AC Q9D700;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 2310043K02RIK PROTEIN.
GN 2310043K02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
```

```

DR EMBL: AK009781; BAB26501.1;
DR MGD: MGI:1914209; 2310043K02RIK.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE.1; 1.
DR PROSITE: PS00056; TYR_PHOSPHATASE.2; 1.
DR PROSITE: PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase.
KW SEQUENCE 198 AA; 22419 MW; 11F0CAF4B5620F0E CRC64;

Query Match 37.9%; Score 440; DB 11; Length 198;
Best Local Similarity 54.3%; Pred. No. 5.3e-34;
Matches 89; Conservative 24; Mismatches 49; Indels 2; Gaps 1;

Qy 39 FELERLFWKSPQYTHVNEVMPKLYIGDEATLDRYLQKAGFTHLVNAAGHGMVNDTGP 98
| | | | | : | : | | | | | : | : | | | | | : | : | | | | |
Db 33 FELERLLYTGKTACHNADEVMPGLYLGQDMANNRRRLRGITHTLVNASHNRW--GTP 90
| | | | | : | : | | | | | : | : | | | | | : | : | | | | |
Qy 99 DYIRDMIOYHGVEADLPTFDLSVFFYPAAAFIDRALSDHSHKILVHCVMGRSRSATLV 158
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Db 91 EAYEGIGIRYLGVEAHDSPAFDMSIHFTQAAFDIHRALSQPGGKTLVHCACVGSRSATLV 150
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Qy 159 LAYLMIHKDMTLVDAIQVAKNRCVLPNRFGLKQLRELDKQLVQ 202
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |
Db 151 LAYLMLYHHLTLVEAIKKVKDHGRIIPNRFGLQLLALDRRLRQ 194
| | | | | : | : | | | | | : | | | | | | | | | | | | | | |

RESULT 6
Q9VWN2 PRELIMINARY; PRT; 245 AA.
ID Q9VWN2
AC Q9VWN2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG7378 PROTEIN.
GN CG7378.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G.C., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beszon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bock J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin Z.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```

DT	01-JUN-2001 (TREMBLREL. 17, Last sequence update)
DE	UNKNOWN (PROTEIN FOR MGC:2627).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;	
RP	[1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=NEUROBLASTOMA;
RA	Strausberg R.;
RL	Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; BC003115; AAH03115.1;
SQ	SEQUENCE 86 AA; 9652 MW; 4DD29AFD5989528B CRC64;
:	
	Query Match 20.1%; Score 233; DB 4; Length 86;
	Best Local Similarity 58.5%; Pred. No. 7.6e-15;
	Matches 48; Conservative 15; Mismatches 19; Indels 0; Gaps
QY	121 LSVFFYPAAFTIDRALSDSHSKILVHCVMGRSRSATLVLAYLIHKDMTLVDIAQQAVKN 180 : :   1 MSIEQTAADEFTHRALSQPGKILVHCACVGVSRSATLVLAYLMLYHHLLTVEAIKKVKDH 60
Dy	
QY	181 RCVLPNRGFLKQLRELDKOLVQ 202   :   61 RGIPNRGELRQLALDLRLRQ 82
Dy	



RESULT 9  
Q9NRW4 PRELIMINARY; PRT; 184 AA.  
AC Q9NRW4;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE X.  
GN MKPX.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gu J., Huang Q., Yu Y., Xu S., Wang Y., Han Z., Chen Z., Zhou J.,  
RA Tu Y., Gu W., Fu G., Huang C.;  
RT "Novel genes expressed in hematopoietic stem/progenitor cells from  
RT Myelodysplastic Syndromes patient."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF165519; AAF86649.1;  
DR InterPro; IPR000340; DS\_phosphatase.  
DR SMART; SM00195; DSPC; 1.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR PFam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Kinase.  
SQ SEQUENCE 184 AA; 20910 MW; B3F962A087C2BA20 CRC64;  
  
Query Match 17.7%; Score 205; DB 4; Length 184;  
Best Local Similarity 35.7%; Pred. No. 9.3e-12;  
Matches 61; Conservative 22; Mismatches 66; Indels 22; Gaps 5;  
  
QY 55 VNEVMPKLYIGDEATDRLRYLQKAGFTHLVNAAGHGNVNDGPDYRDM--DIQVHGVE 112  
DB 5 MNKILPGLYIGNFKDARDAEQLSKNKKVTHLSV-----HDSARPMEGVKYLCP 54  
QY 113 ADDLFTFDLSVFYPAAFIDRALSDHSHKILVCHVCMGRSRTATLVAYLMTHKDMLVD 172  
DB 55 AADSPSNLTRHFESIKFIHCECLRGES-CLVHCLAGVSRVTLVIAVIMTVDFGWED 113  
QY 173 ATQVAKNR-CVLPNRFGLKQLRELDKQLVQRRR-----SQRDGEE 214  
DB 114 ALHTVRAGRSCANPNVGRQLOEFKEHVOYRWLKEEYGESPLQDAEE 164  
  
RESULT 10  
O44128 PRELIMINARY; PRT; 365 AA.  
AC O44128;  
DT 01-JUN-1998 (Tremblrel. 06, Created)  
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE C05B10.1 PROTEIN.  
GN C05B10.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX WATERSTON R.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX WATERSTON R.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF036685; AAB88308.1;  
DR HSSP; Q16828; IMKP.  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR001763; Rhodanese domain.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR PFam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.  
DR SMART; SM00450; RHOD; 1.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
SQ SEQUENCE 365 AA; 41101 MW; IE416C0E9693AF66 CRC64;

Query Match 17.6%; Score 204.5; DB 5; Length 365;  
Best Local Similarity 38.6%; Pred. No. 2.5e-11;  
Matches 54; Conservative 24; Mismatches 55; Indels 7; Gaps 3;  
  
QY 62 LVIGDEATDRLRYLQKAGFTHLVNAAGHGNVNDGPDYRDMIOYHGVEADDLPTFDL 121  
DB 188 LYLGAETAKNRDLKLVKYSISHVINVTSLNPT-----FEEDPNMYRLRISADDNASHNL 242  
QY 122 SVFFYPAAFIDRALSDHSHKILVCHVCMGRSRTATLVAYLMTHKDMLVDATQVAK-N 180  
DB 243 TKFFPEAISFIDARRND-SACLHCLAGISRVTLCLAYLMKTEMCTLDSAYEVWQKRN 301  
QY 181 RCVLPNRFGLKQLRELDKOL 200  
DB 302 ASIAPNFHEMGOLTDYEKML 321  
  
RESULT 11  
Q9ZR37 PRELIMINARY; PRT; 198 AA.  
AC Q9ZR37;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DE DSPTP1 PROTEIN.  
GN DSPTP1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99155801; PubMed=10036776;  
RA Gupta R., Huang Y., Kieber J., Luan S.;  
RT "Identification of a dual-specificity protein phosphatase that  
RT inactivates a MAP kinase from Arabidopsis.";  
RL Plant J. 16:581-589(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gupta R.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y18620; CAA77232.1;

```
DR HSP: Q16828; 1MKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 198 AA; 22017 MW; EBF1C98A177E6450 CRC64;

Query Match
Best Local Similarity 17.5%; Score 203.5; DB 10; Length 198;
Matches 54; Conservative 20; Mismatches 51; Indels 21; Gaps 5;

QY 62 LVIGDEATLDYRLQKAGFTHLNA-----AHGRWNVDTCDDYRDMDIQYHGVEADD 115
DB 58 LVLSGVAASAKNVKLSYNTHTLVASSLRPAH-----PD-----DFYKVVRRVD 104
QY 116 LPTFDLSVFFYPAAAFIDRALSDHSHKLVHCVMGRSRSATLVLAYLMIHKDMTLVDATO 175
DB 105 KEDTNLEMYFDECDVDFIDEAKRQGS-VLVHCFVSKRSVTIVAYLMMKKHGTLAQALQ 163

QY 176 QVAKNRCVL-PNRGFLKQLRELDKQL 200
DB 164 HVKSKRPVSPNAGFIROLQDLEKSM 189

RESULT 12
ID Q9LUG6 PRELIMINARY; PRT; 198 AA.
AC Q9LUG6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE DUAL-SPECIFICITY PROTEIN PHOSPHATASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty P1 and TAC
clones.";
RL DNA Res. 7:131-135(2000).
DR EMBL; AB023036; BAB02780.1; .
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 198 AA; 22113 MW; 815BBCBE3BF696A CRC64;

Query Match
Best Local Similarity 17.5%; Score 203.5; DB 10; Length 198;
Matches 54; Conservative 20; Mismatches 51; Indels 21; Gaps 5;

QY 62 LVIGDEATLDYRLQKAGFTHLNA-----AHGRWNVDTCDDYRDMDIQYHGVEADD 115
DB 58 LVLSGVAASAKNVKLSYNTHTLVASSLRPAH-----PD-----DFYKVVRRVD 104
QY 116 LPTFDLSVFFYPAAAFIDRALSDHSHKLVHCVMGRSRSATLVLAYLMIHKDMTLVDATO 175

Query Match
Best Local Similarity 17.4%; Score 201.5; DB 10; Length 946;
Matches 62; Conservative 24; Mismatches 75; Indels 29; Gaps 7;

QY 27 EGEEDYCTPGAFELERLEFWKSPQYTHVNE-----VMPKLYIGDEATA 70
DB 685 EGSNNST-DAYELK-----VRLEHLERISLISKAANTEKPSMIQENLFIGGLAA 736
QY 71 LDYRLQKAGFTHLNAAGHGRWNVDTCDDYRDMDIQYHGVEADDLPTFDLSVFFYPAA 130
DB 737 RSIYTLQHLGTHVLCLCAN--EIGSDTQYDDL-FEYONFISITDDSDNIESIFEALD 793
QY 131 FIDRALSDHSHKLVHCVMGRSRSATLVLAYLMIHKDMTLVDATIQQVAK-NRCVLPNRGF 189
DB 794 FIKHG-EETGGKILVHCFEGRSRSATVVLAYLMLQKLTLLLEAWSKLKRVHRRAPNDGF 852
QY 190 LKQRELQKQ 199
DB 853 ARILNLDKK 862

RESULT 14
ID Q99N11 PRELIMINARY; PRT; 184 AA.
AC Q99N11;
DT 01-JUN-2001 (Tremblrel. 17, Created)
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:44:11 : Search time 12.75 seconds  
(without alignments)  
388.292 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 220

Sequence: 1 MTSGEVKTSLKNAYSSAKRL.....VQRRRSQRQDGEEDGREL 220

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 212352 seqs, 22503292 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	12	5.5	207	4	US-09-013-881-4
2	12	5.5	263	4	US-09-164-193-5

ALIGNMENTS

RESULT 1

US-09-013-881-4  
: Sequence 4, Application US/09013881  
: Patent No. 6132964

GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/013.881  
FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BILLINGS, LUCY J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0470 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 207 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT02

CLONE: 971204

US-09-013-881-4

Query Match 5.5%; Score 12; DB 4; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.00039;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLAYLM 163

Db 153 SRSATLVLAYLM 164

RESULT 2

US-09-164-193-5

: Sequence 5, Application US/09164193C

: Patent No. 6258582

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as ame

FILE REFERENCE: MNI-051

CURRENT APPLICATION NUMBER: US/09/164,193C

CURRENT FILING DATE: 1998-09-30

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 263

TYPE: PRT

ORGANISM: Homo sapiens

US-09-164-193-5

Query Match 5.5%; Score 12; DB 4; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.00048;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 SRSATLVLAYLM 163

Db 114 SRSATLVLAYLM 125

Search completed: February 8, 2002, 15:46:05  
Job time: 114 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2002, 15:42:06 ; Search time 11.67 seconds  
(without alignments)  
691.196 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEYKTSKLNAYSAKRL.....VOQRRRSQRDGEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	346.5	29.8	185	1	DUS3_HUMAN	P51452	homo sapien
2	196.5	16.9	384	1	DUS9_HUMAN	Q99956	homo sapien
3	192.5	16.6	394	1	DUS4_HUMAN	Q13115	homo sapien
4	192.5	16.6	395	1	DUS4_RAT	Q62767	rattus norv
5	184.5	15.9	318	1	DUS2_MOUSE	Q05922	mus musculus
6	180.5	15.5	314	1	DUS2_HUMAN	Q05923	homo sapien
7	180	15.5	375	1	DUS4_CHICK	Q9PW71	gallus gall
8	169	14.6	367	1	DUS1_HUMAN	P28562	homo sapien
9	168.5	14.5	384	1	DUS5_RAT	O54838	rattus norv
10	166	14.3	280	1	DUS7_RAT	Q63340	rattus norv
11	166	14.3	322	1	DUS7_HUMAN	Q16829	homo sapien
12	166	14.3	367	1	DUS1_MOUSE	P28563	mus musculus
13	166	14.3	367	1	DUS1_RAT	Q64623	rattus norv
14	163	14.0	625	1	DUS8_HUMAN	Q13202	homo sapien
15	163	14.0	663	1	DUS8_MOUSE	O09112	mus musculus
16	161.5	13.9	381	1	DUS6_HUMAN	Q16828	homo sapien
17	161.5	13.9	381	1	DUS6_RAT	Q64346	rattus norv
18	157.5	13.6	384	1	DUS5_HUMAN	Q16690	homo sapien
19	136.5	11.8	276	1	PTP3_CHLEU	Q39491	chlamydomon
20	132.5	11.4	619	1	VHPI1_CAEEL	Q10038	caenorhabdi
21	132	11.4	209	1	Y1L3_YEAST	P40479	saccharomyc
22	128.5	11.1	340	1	DUSC_HUMAN	Q9uni6	homo sapien
23	124.5	10.7	489	1	MSG5_YEAST	P38590	saccharomyc
24	122	10.5	364	1	PVH1_YEAST	Q02256	saccharomyc
25	113	9.7	171	1	VH01_VACCV	P07239	vaccinia vi
26	111	9.6	292	1	Y042_CAEEL	P34680	caenorhabdi
27	108	9.3	171	1	VH01_RACVI	P80994	raccoon pox
28	107	9.2	171	1	VH01_VACCC	P20495	vaccinia vi
29	106	9.1	171	1	VH01_VARV	P33064	variola vir
30	98	8.4	160	1	PTP2_NPVOF	O10273	orgyia pseu
31	97.5	8.4	1442	1	PTPG_MOUSE	Q05909	mus musculus
32	93.5	8.1	1445	1	PTPG_HUMAN	P23470	homo sapien
33	91	7.8	551	1	CC14_YEAST	Q00684	saccharomyc

RESULT 1  
DUS3\_HUMAN  
ID P51452; STANDARD; PRT; 185 AA.  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 3 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE VHR).  
GN DUSP3 OR VHR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93101689; PubMed=1281549;  
RA Ishibashi T., Bottaro D.P., Chan A., Miki T., Aaronson S.A.;  
RT "Expression cloning of a human dual-specificity phosphatase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:12170-12174(1992).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RX MEDLINE=96243129; PubMed=8650341;  
RA Yuwaniyama J., Denu J.M., Dixon J.E., Saper M.A.;  
RT "Crystal structure of the dual specificity protein phosphatase VHR.";  
RL Science 272:1328-1331(1996).  
CC -!- FUNCTION: THIS PROTEIN SHOW BOTH ACTIVITY TOWARD TYROSINE-  
CC PROTEIN PHOSPHATE AS WELL AS WITH SERINE-PROTEIN PHOSPHATE.  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O -  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L05147; AAA35777.1;  
CC PDB; 1VHR; 20-JUN-96.  
CC MIM; 600183;  
CC InterPro; IPR000340; DS\_phosphatase.  
CC InterPro; IPR000387; TYR\_phosphatase.  
CC Pfam; PF00782; DSPC; 1.  
CC SMART; SM00195; DSPC; 1.  
CC PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
CC PROSITE; PS0056; TYR\_PHOSPHATASE\_2; 1.  
CC PROSITE; PS0054; TYR\_PHOSPHATASE\_DUAL; 1.  
CC Hydrolase; 3D-structure.  
CC ACT\_SITE 124 124  
CC SEQUENCE 185 AA; 20478 MW; C1045DD9B226FD94 CRC64;

#### ALIGNMENTS





[illegible]

RESULT 4

DUS4_RAT	
ID	DUS4_RAT
AC	Q62767;
DT	15-JUL-1999 (Rel. 38, Created)
DC	15-JUL-1999 (Rel. 38, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE	(MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE
DE	PHOSPHATASE-2) (MKP-2).
DE	DUSP4 OR MKP2.
GN	Rattus norvegicus (Rat).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxId=10116;
RI	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Phеоchromocytoma;
RC	MEDLINE=95301550; PubMed=7782322;
RX	Misra-Press A., Rim C.S., Yao H., Roberson M.S., Stork P.J.S.;
RT	"A novel mitogen-activated protein kinase phosphatase. Structure,
RT	expression, and regulation.";
RL	J. Biol. Chem. 270:14587-14596(1995).
CC	-1- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
CC	DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
CC	ERK1 AND ERK2 (BY SIMILARITY).
CC	-1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC	PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC	-1- TISSUE SPECIFICITY: EXPRESSED AT MODERATE LEVELS IN NEARLY ALL
CC	TISSUES AND CELLS INCLUDING BRAIN, SPLEEN, AND TESTES WITH THE
CC	HIGHER EXPRESSION IN THE HEART AND LUNG AND LOWER EXPRESSION IN
CC	SKETAL MUSCLE AND KIDNEY. UNDETECTABLE IN LIVER. EXPRESSED IN
CC	MANY AREAS OF THE BRAIN WITH VERY STRONG EXPRESSION IN THE
CC	HIPPOCAMPUS, PIRIFORM CORTEX, AND THE SUPRACHIASMATIC NUCLEUS.
CC	-1- INDUCTION: BY MITOGENS AND BY STRESS.
CC	-1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC	TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collabor
CC	between the Swiss Institute of Bioinformatics and the EMBL outstat
CC	the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no
CC	modified and this statement is not removed. Usage by and for commo
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/annou">http://www.isb-sib.ch/annou</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----

EMBL; U23438; AAC52493.1; -  
DR InterPro: IPR001340; DS\_phosphatase.  
DR InterPro: IPR001763; Rhodanese\_domain.  
DR InterPro: IPR001387; TYR\_phosphatase.  
DR Pfam: PF00782; DSPc; 1.  
DR Pfam: PF00581; Rhodanese; 1.  
DR SMART: SM00195; DSPc; 1.  
DR SMART: SM00450; RHOD; 1.  
DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolase; Nuclear protein.  
FT DOMAIN 47 62 CH2 A DOMAIN.  
FT DOMAIN 139 154 CH2 B DOMAIN.  
FT DOMAIN 198 395 PROTEIN-TYROSINE PHOSPHATASE.  
FT ACT\_SITE 281 281 BY SIMILARITY  
SQ SEQUENCE 395 AA; 43187 MW; A90EFFD378A050FD CRC64;

Query Match 16.6%; Score 192.5; DB 1; Length 395;  
Best Local Similarity 34.2%; Pred. No. 2.5e-11;  
Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps

```

QY 57 EVMPLYIGDEATLDYRLQKAGTFHVLNAAHRWNVDTPDY YRDMDIQYHGVEADDL 116
      | : ||| : | | | | | | | | | | | | | | | | | | | | | | |
Db 199 EILPFLVLSAYHAARRDMLDALGITALLNVS-----SDCPNHF-EGHYQYKCIPVEDN 351

QY 117 PTEDLSVFPPAAAFIDRALSDHSHSKLVHCVMGRSRSLTYLAYLMHKDWTLVDAIQQ 176
      | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 HKADISSWFMEATEYID-AVKDCRGRLVHCQAGISRATICIAYLMMKRVLRLEAEFF 310

QY 177 VAKNRCLV-PNRGLFKQLRELDKOLV 201
      | : | : | | | | | | | | | | | | | | | | | | | | |
Db 311 VKQRSIISFNFSFGQLQFESQVL 336
```

RESULT 5  
DUS2\_MOUSE DUS2\_MOUSE STANDARD; PRT; 318 AA.

ID	DUS2_MOUSE	STANDARD;	PRT;	318 AA.
AC Q05922;	Q60640;			
DT 01-FEB-1994	(Rel. 28, Created)			
DT 01-FEB-1994	(Rel. 28, Last sequence update)			
DE 01-NOV-1997	(Rel. 35, Last annotation update)			
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 2	(EC 3.1.3.48) (EC 3.1.3.16)			
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1).				
DE DSP2 OR PAC1 OR PAC-1.				
OS Mus musculus (Mouse).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX NCBI_TaxID=10090;				
[1]				
RN RP	SEQUENCE FROM N.A.			
RX MEDLINE=93206122;	PubMed=7681221;			
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,				
RA Siebenlist U., Kelly K.;				
RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.";				
RL Science 259:1763-1766(1993).				
RN [2]				
RN RP	SEQUENCE FROM N.A.			
RC STRAIN=129;				
RX MEDLINE=95203877;	PubMed=7896276;			
RA Gerondakis S., Economou C., Grumont R.J.;				
RT "Structure of the gene encoding the murine dual specificity tyrosine-				
RT threonine phosphatase PAC1."				
RL Genomics 24:182-184(1994).				
CC -1- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY				
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES				
CC ERK1 AND ERK2.				
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =				
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.				
CC -1- SUBCELLULAR LOCATION: NUCLEAR.				
CC -1- TISSUE SPECIFICITY: IN HEMATOPOIETIC TISSUES SUCH AS SPLEEN AND THYMUS.				

CC -!- INDUCTION: BY MITOGENS.  
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L11330; AAA19666.1; -  
CC DR EMBL: U09268; AAA85136.1; -  
CC DR HSP: P51452; IVHR.  
CC MGD: MGI:101911; Dusp2.  
CC DR InterPro: IPR000340; DS\_phosphatase.  
CC DR InterPro: IPR001763; Rhodanese\_domain.  
CC DR InterPro: IPR000387; TYR\_phosphatase.  
CC DR Pfam: PF00782; DSPC; 1.  
CC DR Pfam: PF00581; Rhodanese; 1.  
CC DR SMART: SM00195; DSPC; 1.  
CC DR SMART: SM00450; RHOD; 1.  
CC DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
CC DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
CC DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
CC KW Hydrolase; Nuclear protein; Alternative splicing.  
CC FT ACT\_SITE 261 261 BY SIMILARITY.  
CC FT VARSPIC 175 179 GGPVE -> VSSDL (IN SHORT ISOFORM).  
CC FT VARSPIC 180 318 MISSING (IN SHORT ISOFORM).  
CC FT CONFLICT 11 12 CE -> WQ (IN REF. 2).  
CC FT CONFLICT 20 20 A -> V (IN REF. 2).  
CC FT CONFLICT 156 156 P -> A (IN REF. 2).  
CC SQ SEQUENCE 318 AA; 34546 MW; A2006ED0FD27E41D CRC64;

Query Match 15.9%; Score 184.5; DB 1; Length 318;  
Best Local Similarity 28.3%; Pred. No. 1.1e-10;  
Matches 56; Conservative 39; Mismatches 84; Indels 19; Gaps 6;  
QY 10 LKNAYSSAKRLSPKMEEGEDYCTPGAFELER-----LFWKGSFQTHVNEWPKLYI 64  
DB 132 LRGGFKSFQYCPDLCSEAPQAALPPAGAENSNDPRVPYDQGGV-----EILPYLYL 186  
QY 65 GDEATLDRLQKAGFTHVLAHGRNVDTPGYRDMDIQYHGVEADLPTFDLSVF 124  
DB 187 GSCNISSDLQGLQACGITAFLNVS-----ASCPNHEGL-FHYKSIPIVEDNQWVEISAW 239  
QY 125 FYPAAFIDRLSDHSHKILVHCVMGRSRATLVLAHGRNVDTPGYRDMDIQYHGVEADLPTFDLSVF 184  
DB 240 FQEAISFID-SVKNSSGGRVHLVHCQAGISRSATICLAYLIQSHRVRLEDAFDFVKQRRGVI 298  
QY 185 -PNRGFLKQLRELDKQLV 201  
DB 299 SPNFSFMGLQLQLETVL 316  
RESULT 6  
DUS2\_HUMAN  
ID DUS2\_HUMAN STANDARD; PRT; 314 AA.  
AC Q05923;  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1).  
GN DUSP2 OR PAC1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=93206122; PubMed=7681221;  
RA Rohan P., Davis P., Moskaluk C.A., Kearns M., Krutzsch H.,  
RA Siebenlist U., Kelly K.;  
RT "PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.";  
RL Science 259:1763-1766(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96070437; PubMed=7590752;  
RA Yi H., Morton C.C., Weremowicz S., McBride O.W., Kelly K.;  
RA "Genomic organization and chromosomal localization of the DUSP2 gene,  
RT encoding a MAP kinase phosphatase, to human 2p11.2-q11.";  
RL Genomics 28:92-96(1995).  
CC -!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY  
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES  
CC ERK1 AND ERK2.  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- TISSUE SPECIFICITY: IN HEMATOPOIETIC TISSUES.  
CC -!- INDUCTION: BY MITOGENS.  
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L11329; AAA50779.1; -  
CC DR EMBL: U23853; AAA86112.1; -  
CC DR HSP: P51452; IVHR.  
CC MIM: 603068; -  
CC DR InterPro: IPR000340; DS\_phosphatase.  
CC DR InterPro: IPR001763; Rhodanese\_domain.  
CC DR InterPro: IPR000387; TYR\_phosphatase.  
CC DR Pfam: PF00782; DSPC; 1.  
CC DR Pfam: PF00581; Rhodanese; 1.  
CC DR SMART: SM00195; DSPC; 1.  
CC DR SMART: SM00450; RHOD; 1.  
CC DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
CC DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
CC DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
CC KW Hydrolase; Nuclear protein.  
CC FT ACT\_SITE 257 257 BY SIMILARITY.  
CC SQ SEQUENCE 314 AA; 34399 MW; FDD3543C6DE10CA5 CRC64;  
Query Match 15.5%; Score 180.5; DB 1; Length 314;  
Best Local Similarity 32.2%; Pred. No. 2.7e-10;  
Matches 47; Conservative 32; Mismatches 58; Indels 9; Gaps 4;  
QY 57 EVMPLYICDEATLDRLQKAGFTHVLAHGRNVDTPGYRDMDIQYHGVEADL 116  
DB 175 EILPYFLGSCSHSSDLQGLQACGITAFLNVS-----ASCPNHEGL-FHYKSIPIVEDN 227  
QY 117 PTFDLSVFFYPAAAFIDRLSDHSHKILVHCVMGRSRATLVLAHGRNVDTPGYRDMDIQYHGVEADL 176  
DB 228 QMVEISAFWQEAIGFID-WVKNSSGGRVHLVHCQAGISRSATICLAYLIQSHRVRLEDAFDF 286  
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201  
DB 287 VKQRRGVISPFPNFSFMGLQLQLETVL 312  
RESULT 7  
DUS4\_CHICK  
ID DUS4\_CHICK STANDARD; PRT; 375 AA.  
AC Q9PW71;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)

20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE  
DE PHOSPHATASE-2) (MKP-2).  
GN DUSP4 OR MKP2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=White leghorn;  
RX MEDLINE=20379359; PubMed=10918612;  
RA Fu S.-L., Waba A., Vogt P.K.;  
RT "Identification and characterization of genes upregulated in cells  
transformed by v-Jun";  
RL Oncogene 19:3537-3545(2000).  
CC -!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY  
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES  
CC ERK1 AND ERK2 (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATASE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AF167296; AAD46656.1; .  
DR HSSP; Q16828; IMKP.  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR001763; Rhodanese\_domain.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR Pfam; PF00782; DSPC; 1.  
DR Pfam; PF00581; Rhodanese; 1.  
DR SMART; SM00195; DSPC; 1.  
DR SMART; SM00450; RHOD; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS00054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolase; Nuclear protein.  
FT DOMAIN 30 45 CH2 A DOMAIN.  
FT DOMAIN 122 137 CH2 B DOMAIN.  
FT DOMAIN 178 375 PROTEIN-TYROSINE PHOSPHATASE.  
FT ACT\_SITE 261 261 BY SIMILARITY.  
SQ SEQUENCE 375 AA; 41052 MW; 179290D0C2BEEF1 CRC64;  
  
Query Match 15.5%; Score 180; DB 1; Length 375;  
Best Local Similarity 30.1%; Pred. No. 3.9e-10;  
Matches 55; Conservative 35; Mismatches 75; Indels 18; Gaps 6;  
  
QY 20 LSPKMEEGEEDYCTPGAFELERLFWKSPQYTHVNEVMPKLYIGDEATADRYLQKA 79  
DB 151 ISPPSAESLDLGFSCGT----PLNDQGPV-----EILPELYLGSAYHAARRDMLDAL 201  
QY 80 GTHVLNAAHGRWVDTGPDYRRMDIQYHGVEADDLPFDLSVFFYPAAAFIDRLSDD 139  
DB 202 GITALLNVS-----SDCPNHF-EGHYQYKCIPIVDNHNKADISSWFMEAEIYD-SVREC 253  
QY 140 HSKILVHCVMGSRATLVLAYLMTKDMTLVDAIQQVAKNCVL-PNRGFLKQLRELDK 198  
DB 254 CGRLVHCQAGISRSATLCLATLMMKKRVKLEKAFEFVQKRSIIISPNFSGMLQLFES 313  
QY 199 QLV 201  
1::

Db 314 QVL 316  
RESULT 8  
DUSL\_HUMAN  
ID DUSL\_HUMAN STANDARD; PRT; 367 AA.  
AC P28562;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 1 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (MAP KINASE PHOSPHATASE-1) (MKP-1) (PROTEIN-TYROSINE PHOSPHATASE  
DE CL100) (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH1).  
GN DUSP1 OR PTPN10 OR MKP1 OR CL100 OR VH1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Forebrain;  
RX MEDLINE=93024952; PubMed=1406996;  
RA Keyes S.M., Emslie E.A.;  
RT "Oxidative stress and heat shock induce a human gene encoding a  
RT protein-tyrosine phosphatase";  
RL Nature 359:644-647(1992).  
CC -!- FUNCTION: DUAL SPECIFICITY PHOSPHATASE THAT DEPHOSPHORYLATES MAP  
CC KINASE ERK2 ON BOTH THR-183 AND TYR-185.  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -!- INDUCTION: BY OXIDATIVE STRESS AND HEAT SHOCK.  
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X68277; CAA48338.1; .  
DR PIR; S29090; S29090.  
DR HSSP; P51452; 1VHR.  
DR MIM; 600714; .  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR001763; Rhodanese\_domain.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR Pfam; PF00782; DSPC; 1.  
DR Pfam; PF00581; Rhodanese; 1.  
DR SMART; SM00195; DSPC; 1.  
DR SMART; SM00450; RHOD; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS00054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolase; Cell cycle.  
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.  
FT ACT\_SITE 258 258 BY SIMILARITY.  
SQ SEQUENCE 367 AA; 39297 MW; 11BD1D39A9FCD51F CRC64;  
  
Query Match 14.6%; Score 169; DB 1; Length 367;  
Best Local Similarity 30.2%; Pred. No. 4.5e-09;  
Matches 48; Conservative 33; Mismatches 64; Indels 14; Gaps 5;  
  
QY 44 LEWKGSPQYTHVNEVMPKLYIGDEATADRYLQKAGETHVLNAAHGRWVDTGPDYRD 103  
DB 168 LYDQGPV-----EILPELYLGSAYHASKOMLDALGITALINVS-----ANCPNHF-E 215  
QY 104 MDIQYHGVEADDLPFDLSVFFYPAAAFIDRLSDDHSHKILVHCVMGSRATLVLAYLM 163  
DB 216 GHYQYKSIPIVDNHNKADISSWFNEAIDFID-SIKNAGGRVFCQAGISRSATLCLAYLM 274  
1::

```

QY 164 IHKDMTLVDAIQVARNKCVL-PNRGFLKQLRELDKQLV 201
DB 275 RTNRVKLDAFAEFVQRORSIIISPNSFWGQLQFESQVL 313

RESULT 9
DUS5_RAT
ID DUS5_RAT STANDARD; PRT; 384 AA.
AC Q54838;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 5 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP-KINASE PHOSPHATASE CPG21).
GN DUSP5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RX MEDLINE=98364306; PubMed=9699150;
RA Hevroni D., Rattner A., Bundman M., Lederfein D., Gabarah A.,
RA Mangelus M., Silverman M.A., Kedar H., Naor C., Kornuc M., Hanoch T.,
RA Seger R., Theill L.E., Nedivi E., Richter-Levin G., Citri Y.;
RT "Hippocampal plasticity involves extensive gene induction and multiple
RT cellular mechanisms.";
RL J. Mol. Neurosci. 10:75-98(1998).
CC -!- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -!- SIMILARITY: CONTAINS THE CDC25 HOMOLGY DOMAINS 2 A AND B (CH2
CC DOMAINS A AND B).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF013144; AAB94858.1;
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein.
KW DOMAIN 24 39
FT DOMAIN 53 74 CH2 A DOMAIN.
FT DOMAIN 79 82 POLY-GLY.
FT DOMAIN 120 135 CH2 B DOMAIN.
FT DOMAIN 180 384 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 263 BY SIMILARITY.
SQ SEQUENCE 384 AA; 42094 MW; 5644069B8D348700 CRC64;

Query Match 14.5%; Score 168.5; DB 1; Length 384;
Best Local Similarity 31.5%; Pred. No. 5.3e-09;
Matches 46; Conservative 25; Mismatches 66; Indels 9; Gaps 3;

QY 57 EWPVKLYIGDEATLDRYLQKAGFTHLVNAAGRWNVDTGPDYYRDMDIQVHGVEADDL 116
DB 181 EILPFLYLGSAHASKCEFLANLHTALLNVSRRTSEACT-----THLHYKKWIPVEDS 233

RESULT 10
DUS7_RAT
ID DUS7_RAT STANDARD; PRT; 280 AA.
AC Q63340;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 7 (EC 3.1.3.48) (EC 3.1.3.16)
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE MKP-X) (FRAGMENT).
GN DUSP7 OR MKPX.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Neuron;
RX MEDLINE=96224012; PubMed=8626780;
RA Muda M., Boschert U., Dickinson R., Martinou J.C., Martinou I.,
RA Camps M., Schlegel W., Arkinstall S.;
RT "MKP-3, a novel cytosolic protein-tyrosine phosphatase that
RT exemplifies a new class of mitogen-activated protein kinase
RT phosphatase.";
RL J. Biol. Chem. 271:4319-4326(1996).
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94186; CAA63896.1;
DR HSP; P51452; 1VHR.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 192 192 BY SIMILARITY.
SQ SEQUENCE 280 AA; 30668 MW; 45D6F49A2F2BDDDF CRC64;

Query Match 14.3%; Score 166; DB 1; Length 280;
Best Local Similarity 31.2%; Pred. No. 6.2e-09;
Matches 59; Conservative 29; Mismatches 65; Indels 36; Gaps 9;

QY 27 EGGEEDYCTPGAFELERLFWKSP---QYTHVNEVWPKLYIG--DEATLDRYLQKAG 80
DB 27 EGGEEDYCTPGAFELERLFWKSP---QYTHVNEVWPKLYIG--DEATLDRYLQKAG 80

```



```
DR HSP; P51452; 1VHR.
DR MGD; MGI:105120; Ptpn16.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00554; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Cell cycle.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258
FT MUTAGEN 258 258 C->S; LOSS OF ACTIVITY.
SQ SEQUENCE 367 AA; 39369 MW; 50B5F90FEBBD19AB CRC64;

Query Match 14.3%; Score 166; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 8.8e-09;
Matches 48; Conservative 32; Mismatches 65; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVWPKLYIGDEATALDRYLQKAGFTHVLNAAHGRNVDTPGPDYRD 103
Db 168 LYDQGGPV-----EILSFYLGSAVHASKMDLDALGITALINVS-----ANCPNHF-E 215
QY 104 MDIOYHGVEADDLPFDLSVFYPAAFIDRALSDHDKILVHCVMGRSRSATLVLAYLM 163
Db 216 GHYQKSPVVDENHKADISSWFNEAIDFID-SIKDAGRVFVHCQGISRSATLCLAYLM 274
QY 164 IHKDMTLVDIAOQVAKRCVL-PNRGFLKQLRELDKQLV 201
Db 275 RTNRVKLDEAFEFVKQRSSIISPNFSGQLLQFESQVL 313

RESULT 13
DUS1_RAT
ID DUS1_RAT STANDARD; PRT; 367 AA.
AC Q64623;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 1 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MAP KINASE PHOSPHATASE-1) (MKP-1) (PROTEIN-TYROSINE PHOSPHATASE
DE CL100).
GN DUSP1 OR CL100.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Lung;
RA Muda M., Schlegel W., Arkinstall S.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DUAL SPECIFICITY PHOSPHATASE THAT DEPHOSPHORYLATES MAP
CC KINASE ERK2 ON BOTH THR-183 AND TYR-185.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR HSP; P51452; 1VHR.
```

```
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00554; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase; Cell cycle.
FT DOMAIN 175 367 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 258 258 BY SIMILARITY.
SQ SEQUENCE 367 AA; 39541 MW; 5112ADF290499139 CRC64;

Query Match 14.3%; Score 166; DB 1; Length 367;
Best Local Similarity 30.2%; Pred. No. 8.8e-09;
Matches 48; Conservative 32; Mismatches 65; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVWPKLYIGDEATALDRYLQKAGFTHVLNAAHGRNVDTPGPDYRD 103
Db 168 LYDQGGPV-----EILSFYLGSAVHASKMDLDALGITALINVS-----ANCPNHF-E 215
QY 104 MDIOYHGVEADDLPFDLSVFYPAAFIDRALSDHDKILVHCVMGRSRSATLVLAYLM 163
Db 216 GHYQKSPVVDENHKADISSWFNEAIDFID-SIKDAGRVFVHCQGISRSATLCLAYLM 274
QY 164 IHKDMTLVDIAOQVAKRCVL-PNRGFLKQLRELDKQLV 201
Db 275 RTNRVKLDEAFEFVKQRSSIISPNFSGQLLQFESQVL 313

RESULT 14
DUS8_HUMAN
ID DUS8_HUMAN STANDARD; PRT; 625 AA.
AC Q13202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH-5).
GN DUSP8 OR VH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
RT "HVH-5: a protein tyrosine phosphatase abundant in brain that
RT inactivates mitogen-activated protein kinase.";
RL J. Neurochem. 65:1823-1833(1995).
CC -!- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U27193; AAA83151.1; -
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95518; CAA64772.1; -
DR MGD; MGI:106626; Ntpt1.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PRO1217; PRICHEXTENS.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein.
KW DOMAIN 28 43 CH2 A DOMAIN.
FT DOMAIN 117 132 CH2 B DOMAIN.
FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT DOMAIN 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-GLY.
FT DOMAIN 577 600 POLY-SER.
FT DOMAIN 311 552 PRO-RICH.
FT ACT_SITE 246 246 BY SIMILARITY.
SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Query Match 14.0%; Score 163; DB 1; Length 663;
Best Local Similarity 31.8%; Pred. No. 3.7e-08;
Matches 49; Conservative 31; Mismatches 52; Indels 22; Gaps 7;

QY 55 VNEVWPKLYIGDEATLDRIYRLOKAGFTHLVNAHGRNWNVDGPDYRDMDIQYHGVEAD 114
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 161 LTRILPHLYLGSKQDVLNKDLMTQNGISYVLNAS-----NSCPKPDFI--CESRFMRVPIN 214
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 D-----LPTFDLSVFFYPAAAFIDRA-LSDDHSHKILVHCVMGRSRSATLVLAYLMHKD 167
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 215 DNYCEKLLPWLDKSI-----EFIDKAKLSS--CQVIVHCLAGISRSATIAIAYIMKTWG 266

QY 168 MTLVDAIQOQVAKNR-CVLPNRFGLKQLRELDKQL 200
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 MSSDDAYRFVKDRRPSISPNFNLGQLLEYERSL 300

Search completed: February 8, 2002, 15:45:16
Job time: 190 sec

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: February 8, 2002, 15:40:56 ; Search time 15.19 Seconds  
(without alignments)  
1103.252 Million cell updates/sec

Title: US-09-847-519A-2

Perfect score: 1161

Sequence: 1 MTSGEVKTSLKNAYSAKRL.....VOQRRSQRQDEEDGREL 220

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68.\*

1: pirl.\*

2: pirl2.\*

3: pirl3.\*

4: pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346.5	29.8	185	1 A47196	dual specificity p
2	204.5	17.6	365	2 T32494	hypothetical prote
3	192.5	16.6	303	3 T46405	hypothetical prote
4	192.5	16.6	393	2 A56947	dual specificity p
5	192.5	16.6	394	2 A56115	dual specificity p
6	184.5	15.9	314	2 B57126	dual specificity p
7	180.5	15.5	314	1 A57126	dual specificity p
8	174	15.0	223	2 T49365	protein tyrosine p
9	169	14.6	367	1 T39090	dual specificity p
10	168	14.5	205	2 T49364	protein tyrosine p
11	168	14.5	226	2 T21380	hypothetical prote
12	166	14.3	367	1 S24111	dual specificity p
13	166	14.3	367	2 S22655	dual specificity p
14	160.5	13.8	142	2 T03074	dual specificity p
15	159	13.7	283	2 G84458	probable protein p
16	158	13.6	330	2 T39698	protein tyrosine p
17	157.5	13.6	384	1 T38890	dual specificity p
18	150.5	13.0	186	2 T16056	hypothetical prote
19	137	11.8	272	2 T18915	hypothetical prote
20	136.5	11.8	276	2 T48906	protein-tyrosine-p
21	134.5	11.6	204	2 T17802	hypothetical prote
22	132.5	11.4	619	2 T15969	hypothetical prote
23	132	11.4	209	1 S48459	probable dual spec
24	124.5	10.7	489	1 S58725	dual specificity p
25	124	10.7	771	2 T47666	phosphatase-like p
26	122	10.5	364	1 S31304	protein-tyrosine-p
27	113	9.7	171	1 Q0VZHI	dual specificity p
28	111	9.6	292	2 S41012	hypothetical prote
29	110.5	9.5	169	2 T30684	probable dual spec

## ALIGNMENTS

### RESULT 1

A47196

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) DUSP3 [validated] - human  
N:Alternate names: protein-tyrosine-phosphatase VHR; VHL-related dual-specificity ph  
C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence\_revision 01-Mar-1996 #text\_change 15-Sep-2000

C:Accession: A47196; A58760

B:Shibashi, T.; Bottaro, D.P.; Chan, A.; Mikl, T.; Aaronson, S.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 12170-12174, 1992

A:Title: Expression cloning of a human dual-specificity phosphatase.

A:Reference number: A47196; MUID:93101689

A:Accession: A47196

A:Molecule type: mRNA

A:Residues: 1-185 <ISH>

A:Cross-references: GB:L05147; NID:gl81839; PIDN:AAA35777.1; PID:gl81840

A:Experimental source: fibroblasts

A:Note: sequence extracted from NCBI backbone (NCBIN:120790, NCBI:P:120791)

B:Denu, J.M.; Zhou, G.; Wu, L.; Zhao, R.; Yuvaniyama, J.; Saper, M.A.; Dixon, J.E.

J. Biol. Chem. 270, 3796-3803, 1995

A:Title: The purification and characterization of a human dual-specific protein tyro

A:Reference number: A58760; MUID:95181338

A:Accession: A58760

A:Molecule type: protein

A:Residues: 2-7 <DEN>

R:Yuvaniyama, J.; Denu, J.M.; Dixon, J.E.; Saper, M.A.

Submitted to the Brookhaven Protein Data Bank, February 1996

A:Reference number: A66852; PDB:1VHR

A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 8-185

R:Zhou, G.; Denu, J.M.; Wu, L.; Dixon, J.E.

J. Biol. Chem. 269, 28084-28090, 1994

A:Title: The catalytic role of Cys(124) in the dual specificity phosphatase VHR.

A:Reference number: A55447; MUID:95050584

A:Contents: annotation; active site

C:Genetics:

A:Gene: GDB:DUSP3; VHR

A:Cross-references: GDB:342110; OMIM:600183

A:Map position: 17q21-17q21

C:Function:

A:Description: catalyzes hydrolysis of peptidyl-phosphoserine, -phosphothreonine and  
C:Superfamily: dual specificity phosphoprotein phosphatase DUSP3; VHL-type dual spec

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:2-185/Product: dual specificity phosphatase 3 #status predicted <MAT>

F:37-177/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

F:124/Active site: Cys (phosphocysteine intermediate) #status experimental

F:130/Binding site: substrate phosphate (Arg) #status predicted

\* Query Match 29.8%; Score 346.5; DB 1; Length 185;

\* Best Local Similarity 43.8%; Pred. No. 1.2e-24;

\* Matches 78; Conservative 35; Mismatches 54; Indels 11; Gaps 3;

QY 37 GAFE-----LERLFNKGSPQYT-----HVNVEVMPKLYIGDEATALDRYRLQKAGFTHVLA 87

Db 3 GSFELSVQDLNLLSDGSCYSLPQPCNEVTPRIYVGNASVAQDIPKQLKGITHVLNA 62  
QY 88 AHGR--WVWDTPDXYRMDIOYHGVEADLPTFLSLVFFYPAAAFIDRLSDHSHKILV 145  
Db 63 AGRSFMHVNTWANYKDSGITYLGKANDTQEFNLSAYFERAADFIDQALAKNGRVLV 122  
QY 146 HCVMGSRSATLVLAYLMIHKDMLTVDATQVAKNRCVLPNRFGLKQLRELDKQLVQ 203  
Db 123 HCREGYSRPTLVIAIYLMRQKMDYSALSIVRQNRGPNQDGLAQLCOLNDRLAKE 180

## RESULT 2

T32494

hypothetical protein C05B10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T32494

R:Geisler, C.; Wamsley, P.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid C05B10.

A:Reference number: Z21178

A:Accession: T32494

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-365 &lt;GET&gt;

A:Cross-references: EMBL:AF036685; PIDN:AA88308.1; GSPDB:GN00022; CESP:C05B10.1

A:Experimental source: strain Bristol N2; clone C05B10

C:Genetics:

A:Gene: CESP:C05B10.1

A:Map position: 4

A:Intros: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

## Query Match

Best Local Similarity 17.6%; Score 204.5; DB 2; Length 365;

Matches 54; Conservative 24; Mismatches 55; Indels 7; Gaps 3;

QY 62 LVIGDEATLDRLQKAGFTHVLAHAGRWNVDTGPDYRDMDIQYHGVEADLPTFDL 121  
Db 188 LVLGNAETAKNDVLUKKYSIVHNTSNLPT-----FEEDPNMRYLRISADDNASHNL 242  
QY 122 SVFFYPAAAFIDRLSDHSHKILVHCVMGSRSATLVLAYLMIHKDMLTVDATQVAK-N 180  
Db 243 TRFEPETISFIDARRND-SACLVLHCLAGISRSVTICLAYLMKTECTLDSAYENVQRN 301  
QY 181 RCVLPNRFGLKQLRELDKQL 200  
Db 302 ASIAPNFHFMGQUTDYERKML 321

## RESULT 3

T46405

hypothetical protein DKFZp434O1321.1 - human

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 02-Sep-2000

C:Accession: T46405

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23034

A:Accession: T46405

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-303 &lt;AAA&gt;

A:Cross-references: EMBL:AL137704

A:Experimental source: adult testis; clone DKFZp434O1321

C:Genetics:

A:Note: DKFZp434O1321.1

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specificity

## Query Match

Best Local Similarity 16.6%; Score 192.5; DB 2; Length 303;

Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;

Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;  
QY 57 EWPVKLYIGDEATLDRLQKAGFTHVLAHAGRWNVDTGPDYRDMDIQYHGVEADL 116  
Db 107 EILPFLYLGSAVHAARRDMLDALGITALLNVS-----SDCPNHF-EGHYQYKCIPVEDN 159  
QY 117 PTFLDSLVFFYPAAAFIDRLSDHSHKILVHCVMGSRSATLVLAYLMIHKDMLTVDATQ 176  
Db 160 HKADISSWFMEAEIYD-AVKDCRGRLVHVCQAGISRSATICLAYLMKKRVRLEAEFEF 218  
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201  
Db 219 VKQRSIIISPFPFSGQLLQFESQVL 244

## RESULT 4

A56947

dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat

N:Alternate names: mitogen-activated protein kinase phosphatase 2

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 17-Mar-1999

C:Accession: A56947

R:Misra-Press, A.; Rim, C.S.; Yao, H.; Roberson, M.S.; Stork, P.J.S.

J. Biol. Chem. 270, 14587-14596, 1995

A:Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression

A:Reference number: A56947; MUID:95301550

A:Accession: A56947

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-393 &lt;Mis&gt;

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specific

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:202-333/Domain: VHI-type dual specificity phosphoprotein phosphatase homology &lt;VHI

F:279/Active site: Cys (phosphocysteine intermediate) #status predicted

F:285/Binding site: substrate phosphate (Arg) #status predicted

## Query Match

Best Local Similarity 16.6%; Score 192.5; DB 2; Length 393;

Matches 50; Conservative 29; Mismatches 58; Indels 9; Gaps 4;

QY 57 EWPVKLYIGDEATLDRLQKAGFTHVLAHAGRWNVDTGPDYRDMDIQYHGVEADL 116  
Db 197 EILPFLYLGSAVHAARRDMLDALGITALLNVS-----SDCPNHF-EGHYQYKCIPVEDN 249  
QY 117 PTFLDSLVFFYPAAAFIDRLSDHSHKILVHCVMGSRSATLVLAYLMIHKDMLTVDATQ 176  
Db 250 HKADISSWFMEAEIYD-AVKDCRGRLVHVCQAGISRSATICLAYLMKKRVRLEAEFEF 308  
QY 177 VAKNRCVL-PNRGFLKQLRELDKQLV 201  
Db 309 VKQRSIIISPFPFSGQLLQFESQVL 334

## RESULT 5

A56115

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human

N:Alternate names: dual specificity phosphatase HVH2

C:Species: Homo sapiens (man)

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 24-Apr-1998

C:Accession: A56115

R:Guan, K.L.; Butch, E.

J. Biol. Chem. 270, 7197-7203, 1995

A:Title: Isolation and characterization of a novel dual specific phosphatase, HVH2,

A:Reference number: A56115; MUID:95221370

A:Accession: A56115

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-394 &lt;GUA&gt;

A:Cross-references: GB:U01108

C:Genetics:

A:Gene: GDB:DUSP4; HVH2; MKP-2

A:Cross-references: GDB:433893

[illegible]

Db 56 GITHIICI---RQNEANFIKPNF---QQLFRYLVDIADNPVENIIRFPMTKEFDGSL 110  
QY 137 SDDHSHKILVHCVMGSRSSATLVLAIVMIHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLRE 195  
Db 111 QNG-GKVLVHGNAGISRSAAAFVIAVIMETFGMKYRDAFAYVOERRFCINPNAGFVHQLQ 169  
QY 196 LD 197  
Db 170 YE 171

RESULT 9  
S29090

N:Alternate names: protein-tyrosine phosphatase (EC 3.1.3.-) 1 - human  
C:Species: Homo sapiens (man)  
C:Date: 25-Feb-1994 #sequence\_revision 02-May-1994 #text\_change 11-Jun-1999  
C:Accession: S29090; A53052  
R:Keyse, S.M.; Emelie, E.A.  
Nature 359, 644-647, 1992  
A:Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase  
A:Reference number: S29090; MUID:93024952  
A:Accession: S29090  
A:Molecule type: mRNA  
A:Residues: 1-367 <KEY>  
A:CROSS-references: EMBL:X68277; NID:g29980; PIDN:CAA48338.1; PID:g29981  
R:Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.  
J. Biol. Chem. 269, 3596-3604, 1994  
A:Title: Isolation and characterization of a human dual specificity protein-tyrosine phosphatase  
A:Reference number: A53052; MUID:94148864  
A:Accession: A53052  
A:Molecule type: DNA  
A:Residues: 1-367 <KWA>  
A:Experimental source: leukocyte  
A:Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804, NCBIN:143806)  
C:Genetics:  
A:Gene: GDB:DUSP1; PTPN10  
A:CROSS-references: GDB:136197; OMIM:600714  
A:Map position: 5q34-5q34  
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specificity phosphatase 1; heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced phosphatase  
F:181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>  
F:258/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.6%; Score 169; DB 1; Length 367;  
Best Local Similarity 30.2%; Pred. No. 6.4e-08;  
Matches 48; Conservative 33; Mismatches 64; Indels 14; Gaps 5;

QY 44 LFWKSPQYTHVNEVWPVKLYIGDEATLDRLYRQKAGFTHVLNAAHGRWNVDTGPDYRD 103  
Db 168 LYDQGGPV-----EILPFLYLGSAVHASKOMLDALGITALINVS-----ANCPNHF-E 215  
QY 104 MDIOYHGVEADLPFTDLSVFFYPAAAFIDRLSDHSHKILVHCVMGSRSSATLVLAIVM 163  
Db 216 GHYQKSPVDEHNHAKDISSEFNEAIDFID-SIKNAGRVFVHCGAISRATICLAIVM 274  
QY 164 IHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLRELDKQLV 201  
Db 275 RTNRVKLDEAFEVQRRSIISPNFSFMQLLQFESQVL 313

RESULT 10  
I49364

C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49364  
R:Wishart, M.J.; Denu, J.M.; Williams, J.A.; Dixon, J.E.  
J. Biol. Chem. 270, 26782-26785, 1995  
A:Title: A single mutation converts a novel-phosphotyrosine binding domain into a dual-s

A:Reference number: I49364; MUID:96070766  
A:Accession: I49364  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-205 <RES>  
A:CROSS-references: EMBL:U34973; NID:g1063624; PIDN:AAA87036.1; PID:g1063625  
C:Superfamily: VHL-type dual specificity phosphoprotein phosphatase homology  
F:36-174/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

Query Match 14.5%; Score 168; DB 2; Length 205;  
Best Local Similarity 30.2%; Pred. No. 3.8e-08;  
Matches 54; Conservative 30; Mismatches 67; Indels 28; Gaps 7;

QY 22 PMKEEGEEEDYCTPGAFELERLFWKSPQYTHVNEVWPVKLYIGDEATLDRLYR--LQKA 79  
Db 12 PQCKDDAEWTY-----PMRREMQEVLPGFLGYPSSAMSKLPIQKH 55  
QY 80 GFTHVLAHGRWNVDTG---PDYRDMDIQYHGVEADLPFTDLSVFFYPAAAFIDRLAL 136  
Db 56 GITHIICI---RQNEANFIKPNF---QQLFRYLVDIADNPVENIIRFPMTKEFDGSL 110  
QY 137 SDDHSHKILVHCVMGSRSSATLVLAIVMIHKDMTLVDATQOQVAKNR-CVLPNRRGFLKQLR 194  
Db 111 QNG-GKVLVHGNAGISRSAAAFVIAVIMETFGMKYRDAFAYVOERRFCINPNAGFVHQLQ 168

RESULT 11  
T21380

hypothetical protein F26A3.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21380  
R:McMurray, A.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: T19415  
A:Accession: T21380  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-226 <WIL>  
A:CROSS-references: EMBL:Z78419; PIDN:CAB01700.1; GSPDB:GN000019; CESP:F26A3.4  
A:Experimental source: Clone F26A3  
C:Genetics:  
A:Gene: CESP:F26A3.4  
A:Map position: 1  
A:Introns: 117/2; 150/3; 186/3

Query Match 14.5%; Score 168; DB 2; Length 226;  
Best Local Similarity 30.8%; Pred. No. 4.3e-08;  
Matches 56; Conservative 35; Mismatches 73; Indels 18; Gaps 7;

QY 44 LFWKSPQYTHVNEVWPVKLYIGDEATLDRLYRQKAGFTHVLNAAHGRWNVDTGPDYRD 103  
Db 3 LSFVRNPEYAAMSEIVPGLFCG-VSALSDEMKKHKKTHIINAT-----TEVPNLS 54  
QY 104 M-DIOYHGVEADLPFTDLSVFFYPAAAFID---RALSDDHSHKILVHCVMGSRSSATLVL 159  
Db 55 LGDIQRTKLWLEDTP-----QTYIYPHLELQSDIQALIDAGGKVLVHCVAGVRSASICL 110  
QY 160 AVLMHKDMTLVDATQOQV-AKNCVLPNRRGFLKQLRELDKQLVQORRRSQDGEEDGR 218  
Db 111 AFLKYRCNRLREAYHLMKSRWVRPNLGFWRQLIAYE-QNVKENAGSVRLVRDEAQPE 169  
QY 219 EL 220  
Db 170 QL 171

RESULT 12  
S24411

N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor ty



QY 26 EGE---EEDYCTPGAFELERLFWKGSPOYTHV---NEVWPK-LYIGDEATALDRYRQ 77  
Db 20 EGEVCGVCGCHCMPYSSDTV-----AQQVHVSAPPSIILPEFLYLGSDNASRSELLK 73  
QY 78 KAGFTHVLNAHGRMNVDTPDYRDMDIQYHGVEADDLPTDLSVFFYPAAAFIDRALS 137  
Db 74 TQGISRVLN-----TVPWCQNLN-SFTYHGLDNEKVLQFD-----DAIKFLDQC-E 119  
QY 138 DDHSKILVHCVMGRSRSATLVLAYLMIHKDMLVDAIQOVAKNR 181  
Db 120 KDKARVLVHCMGSKSRSPAVVAYLMMKRGWRLAESHOWVKORR 163

Search completed: February 8, 2002, 15:44:07  
Job time: 191 sec